THE ANNALS
of
MATHEMATICAL STATISTICS

(FOUNDED BY H. C. CARVER)

THE OFFICIAL JOURNAL OF THE INSTITUTE OF MATHEMATICAL STATISTICS

VOLUME XI

1940

THE ANNALS OF MATHEMATICAL STATISTICS

The Annals of Mathematical Statistics is published quarterly by the Institute of Mathematical Statistics, Mt. Royal & Guilford Aves., Baltimore, Md. Subscriptions, renewals, orders for back numbers and other business communications should be sent to the Annals of Mathematical Statistics, Mt. Royal & Guilford Aves., Baltimore, Md., or to the Secretary of the Institute of Mathematical Statistics, P. R. Rider, Washington University, St. Louis, Mo.

Manuscripts for publication in the Annals of Mathematical Statistics should be sent to S. S. Wilks, Fine Hall, Princeton, New Jersey. Manuscripts should be typewritten double-spaced with wide margins, and the original copy should be submitted. Footnotes should be reduced to a minimum and whenever possible replaced by a bibliography at the end of the paper; formulae in footnotes should be avoided. Figures, charts, and diagrams should be drawn on plain white paper or tracing cloth in black India ink twice the size they are to be printed. Authors are requested to keep in mind typographical difficulties of complicated mathematical formulae.

Authors will ordinarily receive only galley proofs. Fifty reprints without covers will be furnished free. Additional reprints and covers furnished at cost.

The subscription price for the Annals is $4.00 per year. Single copies $1.25. Back numbers are available at the following rates:

Vols. I-IV $5.00 each. Single numbers $1.50.
Vols. V to date $4.00 each. Single numbers $1.25.

Photo-Lithoprint Reproduction
EDWARDS BROTHERS, INC.
Lithoprinters
ANN ARBOR, MICHIGAN
1947

INC.
AN

THE ANNALS
of
MATHEMATICAL STATISTICS

(FOUNDED BY H. C. CARVER)

THE OFFICIAL JOURNAL OF THE INSTITUTE OF MATHEMATICAL STATISTICS

Contents

Vol. XI, No. 1 — March, 1940

THE ANNALS OF MATHEMATICAL STATISTICS

Manuscripts for publication in the ANNALS OF MATHEMATICAL STATISTICS should be sent to S. S. Wilks, Fine Hall, Princeton, New Jersey. Manuscripts should be typewritten double-spaced with wide margins, and the original copy should be submitted. Footnotes should be reduced to a minimum and whenever possible replaced by a bibliography at the end of the paper; formulae in footnotes should be avoided. Figures, charts, and diagrams should be drawn on plain white paper or tracing cloth in black India ink twice the size they are to be printed. Authors are requested to keep in mind typographical difficulties of complicated mathematical formulae.

Authors will ordinarily receive only galley proofs. Fifty reprints without covers will be furnished free. Additional reprints and covers furnished at cost.

The subscription price for the ANNALS is $4.00 per year. Single copies $1.25. Back numbers are available at the following rates:

Vols. I-IV $5.00 each. Single numbers $1.50.
Vols. V to date $4.00 each. Single numbers $1.25.

The ANNALS OF MATHEMATICAL STATISTICS is published quarterly by the Institute of Mathematical Statistics. Subscriptions, renewals, orders for back numbers and other business communications should be sent to the Secretary of the Institute of Mathematical Statistics, P. R. Rider, Washington University, St. Louis, Mo.

COMPOSED AND PRINTED AT THE
WAVERLY PRESS, INC.
BALTIMORE, MD., U. S. A.

ON THE UNBIASED CHARACTER OF LIKELIHOOD-RATIO TESTS FOR INDEPENDENCE IN NORMAL SYSTEMS

By Joseph F. Daly

1. Introduction. In the statistical interpretation of experimental data, the basic assumption is, of course, that we are dealing with a sample from a statistical population, the elements of which are characterized by the values of a number of random variables $x^1, \cdots, x^k$. But in many cases we are in a position to assume even more, namely, that the population has an elementary probability law $f(x^1, \cdots, x^k; \theta_1, \cdots, \theta_h)$, where the functional form of $f(x, \theta)$ is definitely specified, although the parameters $\theta_1, \cdots, \theta_h$ are to be left free for the moment to have values corresponding to any point of a set Ω in an h-dimensional space.

Under this assumption, the problem of obtaining from the data further information about the hypothetical distribution law $f(x, \theta)$ is considerably simplified. For it is then equivalent to that of deciding whether or not the data support the hypothesis that the population values of the θ's correspond to a point in a certain subset ω of Ω. For example, we may have reason to believe that the population K has a distribution law of the form

$$f(x^1, x^2; a^1, a^2, A_{11}, A_{12}, A_{22}) = \frac{|A_{ij}|^{\frac{1}{2}}}{2\pi} e^{-\frac{1}{2}\sum_{i,j} A_{ij}(x^i - a^i)(x^j - a^j)}$$

Here the set Ω is composed of all parameter points $(a^1, \cdots, A_{22})$ for which the matrix $\| A_{ij} \|$ $(i, j = 1, 2)$ is positive definite and for which $-\infty < a^i < \infty$. We may wish to decide, on the basis of N independent observations (x^1_α, x^2_α) drawn from K, whether A_{12} has the value zero for the population in question, without concerning ourselves at all about the values of the remaining parameters; in other words, we may wish to test the hypothesis H that the parameter point corresponding to K lies in that subset of Ω for which $A_{12} = 0$. One way to test this hypothesis is to select some (measurable) function $g(x)$ whose value can be determined from the data, say

$$g(x) = \frac{\sum_{\alpha=1}^{N} (x^1_\alpha - \bar{x}^1)(x^2_\alpha - \bar{x}^2)}{\left[\sum_{\alpha=1}^{N} (x^1_\alpha - \bar{x}^1)^2\right]^{\frac{1}{2}}\left[\sum_{\alpha=1}^{N} (x^2_\alpha - \bar{x}^2)^2\right]^{\frac{1}{2}}}$$

Now $g(x)$ is itself a random variable, so that it has a distribution law of its own when its constituent x's are drawn from any particular population K. Suppose then we choose a set of values of $g(x)$, say S, such that the probability is only .05 that $g(x)$ will lie in the set S when the x's are drawn independently from a population K for which the above hypothesis H is true. Ordinarily we would

take S to be of the form $| g(x) | \geq g_0$, and the test would then reject H at the .05 probability level if the computed value of $g(x)$ came out too large. But for all that has been said so far, we are perfectly free to choose a different critical region S, and even a different function $g(x)$. The essential elements of this type of test are then a critical region S, a function of the data g, and a probability level ϵ, such that the probability is $\epsilon = .05$, say, that $g \subset S$ when H is true; in employing the test we reject H at the given probability level whenever the sample value of g falls in the critical region.

By the very nature of the problem, any inferences we make from a sample are subject to possible error. In the kind of test under consideration, the only error we can commit, strictly speaking, is that of rejecting H when it is true (an error of Type I in the terminology of Neyman and Pearson [9]). The risk of such an error is thus known in advance; for if we use the test consistently at, say, the .05 level, we know that the probability is .05 that we shall be led to reject a given hypothesis when it is true. On the other hand, it is quite conceivable that the test may be even less likely to reject H when it is false, or more precisely, when the true θ's correspond to a point of Ω which is not in ω. In this event the test is said to be biased. Let us make this term more definite by proposing the following definitions:

Definition I. *A test is said to be completely unbiased if it has the property that for any probability level* ϵ $(0 < \epsilon < 1)$ *the probability of rejecting H is greater when the θ's correspond to a point of* $\Omega - \omega$ *than when they correspond to a point of* ω.

Definition II. *A test is said to be locally unbiased if the set* Ω *contains a neighborhood U of* ω *such that for any probability level* ϵ $(0 < \epsilon < 1)$ *the probability of rejecting H is greater when the parameter values correspond to a point of* $U - \omega$ *than when they correspond to a point of* ω.

It is the purpose of this paper to consider the question of bias in connection with the Neyman-Pearson method of likelihood ratios [8] as applied to the testing of what may well be called hypotheses of independence in multivariate normal populations. The likelihood ratio method is undoubtedly a very familiar one, since the vast majority of tests in present statistical practice are based on this method. But for the sake of completeness we shall outline it briefly. Let the distribution law of the population K be of the form $f(x^1, \cdots, x^k; \theta_1, \cdots, \theta_h)$ where the θ's may correspond to any point in a set Ω, and let the hypothesis H to be tested be that the θ's actually belong to the subset ω of Ω. Form the likelihood function

$$P_N(x; \theta) = \prod_{\alpha=1}^{N} f(x_\alpha^1, \cdots, x_\alpha^k; \theta_1, \cdots, \theta_h)$$

i.e., the elementary probability law of a sample of N elements drawn independently from K. Denote by $P_N^\Omega(x)$ the maximum of P_N for fixed x where the θ's are allowed to range over Ω; and denote by $P_N^\omega(x)$ the corresponding maximum value when the θ's are restricted to ω. The test criterion is then

$$\lambda = \frac{P^\omega(x)}{P^\Omega(x)}.$$

Evidently λ depends only on the observable quantities x^i_α, and has the range $0 \leq \lambda \leq 1$, with a definite probability law depending on that of the basic population K. In this method the critical region S is taken to be $0 \leq \lambda \leq \lambda_\epsilon$, where λ_ϵ is so chosen that the probability $P\{\lambda \leq \lambda_\epsilon\}$ is ϵ when the parameters of K correspond to a point in ω. (It may be noted here that in all the cases with which we shall have to deal the probability that λ lies in S when H is true is independent of the particular values of the θ's as long as they correspond to a point of ω.) The reason for taking the critical region to be of the form $0 \leq \lambda \leq \lambda_\epsilon$ and not, say, $\lambda'_\epsilon \leq \lambda \leq \lambda''_\epsilon$ or $\bar{\lambda}_\epsilon \leq \lambda \leq 1$ may become clearer when we examine the resulting tests for bias.

The recent work of Neyman and Pearson [10] has led them to lay considerable stress on the importance of unbiased tests. And though their attention has been directed mainly to the broader outlines of the theory of testing hypotheses, they have stimulated other writers to study particular tests of great practical importance. P. C. Tang [11] has obtained the general sampling distribution of $1 - \lambda^{2/N}$ for what we shall call the regression problem with one dependent variate, and has given tables for $P\{\lambda \leq \lambda_\epsilon\}$—essentially proving the unbiased character of the test—which should be extremely useful. His article also contains an excellent discussion of the manner in which this test is related to the well known tests of linear hypotheses [7] and to the ordinary analysis of variance. P. L. Hsu [6] has shown that this same distribution is fundamental in the study of Hotelling's generalized T test [5] (a special but important case of what we shall call the general regression problem), and has proved that (locally) this test is not only unbiased but "most powerful" in a certain sense. On the other hand, it is not true that all likelihood ratio tests are unbiased [2]. Consequently, the knowledge that in a rather wide class of problems which arise in normal sampling theory the method of likelihood ratios furnishes tests which are either locally or completely unbiased would seem to be of some value, even when the exact sampling distribution of the criterion is too complicated to tabulate.

2. **The regression problem with one dependent variate.** Suppose that y is known to be normally distributed about a linear function of the fixed variables $x^1, \cdots, x^r$, so that the family of populations under consideration is characterized by a distribution function of the form

$$f(y \mid x, b, \sigma^2) = (2\pi\sigma^2)^{-\frac{1}{2}} e^{-\frac{1}{2\sigma^2}\left(y - \sum_{i=1}^{r} b_i x^i\right)^2}, \tag{2.1}$$

where the set of admissible values of σ^2 and the b's is

$$\Omega\colon 0 < \sigma^2 < \infty, \qquad -\infty < b_i < \infty.$$

Let H be the hypothesis that the point $(\sigma^2, b_1, \cdots, b_r)$ lies in the subset of Ω defined by

$$\omega\colon b_{q+1} = b_{q+2} = \cdots = b_r = 0.$$

The likelihood ratio appropriate to testing the hypothesis H on the basis of N ($N > r$) independent observations drawn from such a population is then

$$\lambda = \left\{ \frac{\sum_{\alpha=1}^{N} \left(y_\alpha - \sum_{i=1}^{r} \hat{b}_i x_\alpha^i \right)^2}{\sum_{\alpha=1}^{N} \left(y_\alpha - \sum_{k=1}^{q} \hat{b}_k^0 x_\alpha^k \right)^2} \right\}^{\frac{1}{2}N},$$

with the understanding that the values of the fixed variables $x_\alpha^1, \cdots, x_\alpha^r$ associated with the α-th observation have been so chosen that the matrix $\| a^{ij} \| = \left\| \sum_{\alpha=1}^{N} x_\alpha^i x_\alpha^j \right\|$ is positive definite. (The expression in the numerator is the minimum of $\sum_{\alpha=1}^{N} \left(y_\alpha - \sum_{i=1}^{r} b_i x_\alpha^i \right)^2$ for variations of the b's over Ω, while the denominator contains the corresponding minimum for variations of the b's over ω).

In order to show that the test is unbiased, we shall make use of the exact sampling distribution of the quantity

$$\xi = 1 - \lambda^{2/N},$$

first published by P. C. Tang [11]. Writing $\| A_{gh} \|$ for the inverse of the matrix $\| a^{gh} \|$ composed of the first q rows and columns of $\| a^{ij} \|$, let us put

$$G = \frac{1}{2\sigma^2} \sum_{k,l=q+1}^{r} \left(a^{kl} - \sum_{g,h=1}^{q} a^k A_{gh} a^{hl} \right) b_k b_l .$$

Since the critical region $0 \leq \lambda \leq \lambda_\epsilon$ corresponds to the region $1 - \lambda_\epsilon^{2/N} = \xi_\epsilon \leq \xi \leq 1$, it can then be shown that the probability of rejecting H when the population parameters have specified values $\sigma^2, b^1, \cdots, b^r$ is expressed by the series

$$I(G, \xi_\epsilon) = e^{-G} \sum_{\nu=0}^{\infty} \frac{G^\nu}{\nu!} \int_{\xi_\epsilon}^{1} \frac{\xi^{\frac{1}{2}(r-q)+\nu-1}(1-\xi)^{\frac{1}{2}(N-r)-1}}{\mathrm{B}[\frac{1}{2}(r-q)+\nu, \frac{1}{2}(N-r)]} d\xi, \tag{2.2}$$

where

$$\mathrm{B}(u, v) = \frac{\Gamma(u)\Gamma(v)}{\Gamma(u+v)} = \int_0^1 z^{u-1}(1-z)^{v-1} dz.$$

Now G is a positive definite quadratic form in the parameters $b^{q+1}, \cdots, b^r$, so that it vanishes if and only if the hypothesis is true. And if $0 < \epsilon < 1$, then $I(G, \xi_\epsilon)$ is a monotone increasing function of G. For by differentiating (2.2) we obtain

$$\begin{aligned} \frac{\partial}{\partial G} I(G, \xi_\epsilon) = e^{-G} \sum_{\nu=0}^{\infty} \frac{G^\nu}{\nu!} \int_{\xi_\epsilon}^{1} \Bigg\{ & \frac{\xi^{\frac{1}{2}(r-q)+\nu}(1-\xi)^{\frac{1}{2}(N-r)-1}}{\mathrm{B}[\frac{1}{2}(r-q)+\nu+1, \frac{1}{2}(N-r)]} \\ & - \frac{\xi^{\frac{1}{2}(r-q)+\nu-1}(1-\xi)^{\frac{1}{2}(N-r)-1}}{\mathrm{B}[\frac{1}{2}(r-q)+\nu, \frac{1}{2}(N-r)]} \Bigg\} d\xi. \end{aligned} \tag{2.3}$$

And from a property of incomplete Beta functions, which we shall demonstrate in the next section, it follows that each term in the series (2.3) is positive. Accordingly we have

THEOREM I. *The likelihood ratio test for the hypothesis that in a population of type* (2.1) *certain of the regression coefficients are zero, i.e., the hypothesis that y is independent of the fixed variables* $x^{q+1}, \cdots, x^r$, *is completely unbiased.*

Wilks [15] has noted that the ordinary analysis of variance and covariance amounts essentially to testing hypotheses of this nature by means of the function

$$\zeta = \frac{1 - \lambda^{2/N}}{\lambda^{2/N}}.$$

Consequently such tests are also completely unbiased, since the region of rejection is then taken to be of the form $\zeta \geq \zeta_\epsilon$.

3. **An inequality relating to incomplete Beta functions.** Let us write

$$\mathrm{B}(u, v; t) = \int_t^1 z^{u-1}(1 - z)^{v-1}\, dz \qquad (0 \leq t \leq 1).$$

Now,

$$\int_t^1 z^{u-1}(1 - z)^v\, dz = \left.\frac{z^u(1 - z)^v}{u}\right]_t^1 + \frac{v}{u}\int_t^1 z^u(1 - z)^{v-1}\, dz.$$

The integrated term on the right is non-positive, so that

$$\mathrm{B}(u, v + 1; t) \leq \frac{v}{u}\, \mathrm{B}(u + 1, v; t) \tag{3.1}$$

in which the equality holds if and only if $t = 0$ or $t = 1$. Again, since

$$z^u(1 - z)^{v-1} + z^{u-1}(1 - z)^v \equiv z^{u-1}(1 - z)^{v-1},$$

we have

$$\mathrm{B}(u + 1, v; t) + \mathrm{B}(u, v + 1; t) \equiv \mathrm{B}(u, v; t). \tag{3.2}$$

Combining these results, we find that

$$\frac{u + v}{u}\, \mathrm{B}(u + 1, v; t) \geq \mathrm{B}(u, v; t) \tag{3.3}$$

with equality only when $t = 0$ or $t = 1$. Hence we have

LEMMA 1: *If* $0 < t < 1$, *then*

$$\frac{\mathrm{B}(u + 1, v; t)}{\mathrm{B}(u + 1, v)} > \frac{\mathrm{B}(u, v; t)}{\mathrm{B}(u, v)}.$$

4. **The multiple correlation coefficient.** Suppose the distribution law of the underlying population is known to be of the form

$$f(x^1, \cdots, x^t \mid x^{t+1}, \cdots, x^m) = \frac{|B_{ij}|^{\frac{1}{2}}}{\pi^{\frac{1}{2}t}}\, e^{-B_{ij}(x^i - a^i - C^i_p x^p)(x^j - a^j - C^j_q x^q)}. \tag{4.1}$$

The indices appearing in this expression take the values $i, j = 1, \cdots, t$ and $p, q = t + 1, \cdots, m$. The summation convention of repeated indices will be

used, for example, $\sum_{p=t+1}^{m} C^i_p x^p$ will be denoted by $C^i_p x^p$. We shall also have occasion to use indices r, s with the range $r, s = 1, \cdots, m$. The set of possible values of the a's, B's, and C's is

$$\Omega\colon \; \| B_{ij} \| \text{ positive definite}; -\infty < a^i < \infty; -\infty < C^i_p < \infty.$$

We shall consider the λ test for the hypothesis H that x^1 is independent of the remaining variables $x^2, \cdots, x^m$, i.e., that the parameters belong to that subset of Ω defined by

$$\omega\colon \; B_{1k} = 0, \qquad (k = 2, \cdots, t); \qquad C^1_p = 0.$$

Let us write $v^{rs} = \sum_{\alpha=1}^{N} (x^r_\alpha - \bar{x}^r)(x^s_\alpha - \bar{x}^s)$, and assume that the values of the fixed variables x^p_α have been so selected that the matrix $\| v^{pq} \|$ is positive definite. The likelihood ratio can then be expressed in the form

$$\lambda = \left(\frac{| v^{rs} |}{v^{11} \cdot \bar{v}_{11}} \right)^{\frac{1}{2}N} = (1 - R^2)^{\frac{1}{2}N},$$

where $\bar{v}_{11}$ is the complement of v^{11} in the determinant $| v^{rs} |$. If $N \geq m + 1$, the general sampling distribution of R^2 (the multiple correlation coefficient between x^1 and $m - 1$ other variates), for this case in which $x^2, \cdots, x^t$ are subject to sampling variation and the remainder are fixed, is

$$\begin{aligned} F(R^2)\, d(R^2) &= \frac{(1 - \rho^2)^{\frac{1}{2}(N-1)} e^{-\frac{1}{2}y^2} (1 - R^2)^{\frac{1}{2}(N-m)-1} (R^2)^{\frac{1}{2}(m-1)-1}}{\Gamma[\frac{1}{2}(N - m)]} \\ &\times \sum_{\mu=0}^{\infty} \sum_{\nu=0}^{\infty} \frac{\frac{1}{2}(y^2)^\mu (1 - \rho^2)^\mu (\rho^2)^\nu (R^2)^{\mu+\nu} \Gamma^2[\frac{1}{2}(N - 1) + \mu + \nu]}{\mu!\,\nu!\,\Gamma[\frac{1}{2}(N - 1) + \mu]\Gamma[\frac{1}{2}(m - 1) + \mu + \nu]} d(R^2), \end{aligned} \tag{4.2}$$

where

$$1 - \rho^2 = \frac{| B_{ij} |}{B_{11} B^{11}}, \qquad \tfrac{1}{2}y^2 = \frac{v^{pq}}{B^{11}} C^1_p C^1_q, \qquad \| B^{ij} \| = \| B_{ij} \|^{-1}.$$

This distribution was first obtained by Wilks [13], although Fisher [3] had previously treated the two extreme cases in which (1) all independent variables are subject to sampling fluctuation, and (2) all independent variables are fixed.

To simplify the presentation, let us put $\bar{\rho} = \rho^2$, $\bar{y} = \frac{1}{2}y^2$ and $\bar{R} = R^2$, and note that $\bar{y} = 0$ if and only if $C^1_p = 0$ ($p = t + 1, \cdots, m$) while $\bar{\rho} = 0$ if and only if $B_{1k} = 0$ ($k = 2, \cdots, t$), so that $\bar{y} = \bar{\rho} = 0$ means that the hypothesis H is true. On any alternative hypothesis, one or the other or both of these quantities will be positive. Let the region of rejection be taken to be

$$R_\epsilon \leq \bar{R} \leq 1,$$

which corresponds to

$$0 \leq \lambda \leq (1 - R_\epsilon)^{\frac{1}{2}N}.$$

The probability of rejecting H is then

$$
\begin{aligned}
(4.4)\quad I(\bar\rho, \bar y, \bar R_\epsilon) &= e^{-\bar y} \sum_{\mu=0}^{\infty} \sum_{\nu=0}^{\infty} \frac{\bar y^\mu}{\mu!} (1-\bar\rho)^{\frac{1}{2}(N-1)+\mu} \frac{\bar\rho^\nu}{\nu!} \frac{\Gamma[\frac{1}{2}(N-1)+\mu+\nu]}{\Gamma[\frac{1}{2}(N-1)+\mu]} \\
&\quad \times \int_{\bar R_\epsilon}^{1} \frac{\bar R^{\frac{1}{2}(m-1)+\mu+\nu-1}(1-\bar R)^{\frac{1}{2}(N-m)-1}}{B[\frac{1}{2}(m-1)+\mu+\nu, \frac{1}{2}(N-m)]} d\bar R.
\end{aligned}
$$

We shall show that $I(\bar\rho, \bar y, \bar R_\epsilon)$ is a strictly monotone increasing function of $\bar\rho$ for each $\bar y$, and that $I(0, \bar y, R_\epsilon)$ is a strictly monotone increasing function of $\bar y$. First consider $\dfrac{\partial I}{\partial \bar\rho}$. We can write (4.4) in the form

$$
I(\bar\rho, \bar y, R_\epsilon) = e^{-\bar y} \sum_{\mu=0}^{\infty} \frac{\bar y^\mu}{\mu!} \frac{1}{\Gamma[\frac{1}{2}(N-1)+\mu]} \cdot \sum_{\nu=0}^{\infty} \frac{\bar\rho^\nu}{\nu!} (1-\bar\rho)^{\frac{1}{2}(N-1)+\mu} \varphi_{\mu,\nu},
$$

where

$$
\varphi_{\mu,\nu} = \Gamma[\tfrac{1}{2}(N-1)+\mu+\nu] \frac{B[\frac{1}{2}(m-1)+\mu+\nu, \frac{1}{2}(N-m); \bar R_\epsilon]}{B[\frac{1}{2}(m-1)+\mu+\nu, \frac{1}{2}(N-m)]}.
$$

Then, formally,

$$
\begin{aligned}
&\frac{\partial}{\partial\bar\rho}\left(\sum_{\nu=0}^{\infty} \frac{\bar\rho^\nu}{\nu!}(1-\bar\rho)^{\frac{1}{2}(N-1)+\mu}\varphi_{\mu,\nu}\right) \\
&\qquad = \sum_{\nu=0}^{\infty} \frac{\nu\bar\rho^{\nu-1}}{\nu!}(1-\bar\rho)^{\frac{1}{2}(N-1)+\mu} - \sum_{\nu=0}^{\infty} \frac{\bar\rho^\nu}{\nu!}(1-\bar\rho)^{\frac{1}{2}(N-1)+\mu-1}[\tfrac{1}{2}(N-1)+\mu]\varphi_{\mu,\nu}.
\end{aligned}
$$

Taking out the factor $(1-\bar\rho)^{\frac{1}{2}(N-1)+\mu-1}$, we have left

$$
\begin{aligned}
&\sum_{\nu=0}^{\infty} \frac{\nu\rho^{\nu-1}}{\nu!}\varphi_{\mu,\nu} - \sum_{\nu=0}^{\infty} \frac{\nu\rho^{\nu}}{\nu!}\varphi_{\mu,\nu} - \sum_{\nu=0}^{\infty} \frac{\rho^\nu}{\nu!}[\tfrac{1}{2}(N-1)+\mu]\varphi_{\mu,\nu} \\
&\qquad = \sum_{\nu=0}^{\infty} \frac{\rho^\nu}{\nu!}\{\varphi_{\mu,\nu+1} - [\tfrac{1}{2}(N-1)+\mu+\nu]\varphi_{\mu,\nu}\}.
\end{aligned}
$$

And the expression $\varphi_{\mu,\nu+1} - [\frac{1}{2}(N-1)+\mu+\nu]\varphi_{\mu,\nu}$ is the same as

$$
\begin{aligned}
\Gamma[\tfrac{1}{2}(N-1)+\mu+\nu+1]&\left\{\frac{B[\frac{1}{2}(m-1)+\mu+\nu+1, \frac{1}{2}(N-m), \bar R_\epsilon]}{B[\frac{1}{2}(m-1)+\mu+\nu+1, \frac{1}{2}(N-m)]}\right. \\
&\qquad \left. - \frac{B[\frac{1}{2}(m-1)+\mu+\nu, \frac{1}{2}(N-m), \bar R_\epsilon]}{B[\frac{1}{2}(m-1)+\mu+\nu, \frac{1}{2}(N-m)]}\right\}
\end{aligned}
$$

and is therefore positive, by Lemma 1. Consequently

$$
\frac{\partial}{\partial\bar\rho} I(\bar\rho, \bar y, \bar R_\epsilon) \geq 0,
$$

with equality holding only if $\bar\rho = 1$, or if the critical region is taken as the whole interval or the null set.

We have yet to investigate $\frac{\partial}{\partial \bar{y}} I(0, \bar{y}, \bar{R}_\epsilon)$. In this case (4.4) becomes

$$I(0, \bar{y}, \bar{R}_\epsilon) = e^{-\bar{y}} \sum_{\mu=0}^{\infty} \frac{\bar{y}^\mu}{\mu!} \frac{B[\frac{1}{2}(m-1)+\mu, \frac{1}{2}(N-m), \bar{R}_\epsilon]}{B[\frac{1}{2}(m-1)+\mu, \frac{1}{2}(N-m)]}. \tag{4.5}$$

(Note that this agrees with (2.2) if we make use of the relations $r = m, q = 1$, and $B^{11} = 2\sigma^2$.) We then obtain

$$\frac{\partial}{\partial \bar{y}} I(0, \bar{y}, \bar{R}_\epsilon) = e^{-\bar{y}} \sum_{\mu=0}^{\infty} \frac{y^\mu}{\mu!} \left\{ \frac{B[\frac{1}{2}(m-1)+\mu+1, \frac{1}{2}(N-m); \bar{R}_\epsilon]}{B[\frac{1}{2}(m-1)+\mu+1, \frac{1}{2}(N-m)]} - \frac{B[\frac{1}{2}(m-1)+\mu, \frac{1}{2}(N-m); \bar{R}_\epsilon]}{B[\frac{1}{2}(m-1)+\mu, \frac{1}{2}(N-m)]} \right\}$$

which the lemma shows to be positive when $0 < R_\epsilon < 1$.

This concludes the proof of

THEOREM II. *If the underlying population has a distribution law of the form* (4.1), *then the likelihood ratio test for the hypothesis that x^1 is independent of $x^2, \cdots, x^m$, where $x^{t+1}, \cdots, x^m$ are fixed and $x^2, \cdots, x^t$ are subject to sampling variation, is completely unbiased.*

5. **Mutual independence of several sets of random variables.**[1] Let the distribution law of the m-variate population be of the form

$$\frac{|B_{ij}|^{\frac{1}{2}}}{\pi^{\frac{1}{2}m}} e^{-B_{ij}(x^i-a^i)(x^j-a^j)}. \tag{5.1}$$

Here Ω is the set $\| B_{ij} \|$ positive definite; $-\infty < a^i < \infty$. Suppose we wish to test the hypothesis H_I that the variates $\{x^1, \cdots, x^{m_1}\}, \cdots, \{x^{m_{p-1}+1}, \cdots, x^{m_p}\}$ are mutually independent in sets [14], where $0 = m_0 < m_1 < \cdots < m_p = m$. Then the ω set is that defined by

$$\| B_{ij} \| = \| B_{i_1 j_1} \| \dotplus \cdots \dotplus \| B_{i_p j_p} \| = \| B_1 \| \dotplus \cdots \dotplus \| B_p \|,$$

that is, we have $B_{ij} = 0$ unless the indices i and j both relate to the same set of variates.

Associated with the population of random samples O_N ($N \geq m + 1$) drawn from a universe characterized by (5.1), we have the distribution function

$$P(x; B, a) = \frac{|B_{ij}|^{\frac{1}{2}N}}{\pi^{\frac{1}{2}Nm}} e^{-\sum_{\alpha=1}^{N} B_{ij}(x_\alpha^i - a^i)(x_\alpha^j - a^j)}$$

The maximum of P with respect to variations of the parameters B_{ij}, a^i in Ω is

$$P_\Omega = |v^{ij}|^{-\frac{1}{2}N} \left(\frac{N}{2\pi}\right)^{\frac{1}{2}Nm} e^{-\frac{1}{2}N},$$

[1] In this and in subsequent sections an index occurring both above and below indicates summation in accordance with the usual convention.

where

$$v^{ij} = \sum_{\alpha=1}^{N} (x_\alpha^i - \bar{x}^i)(x_\alpha^j - \bar{x}^j).$$

And the maximum when the parameters are restricted to ω is

$$P_\omega = [v_1 \cdots v_p]^{-\frac{1}{2}N} \left(\frac{N}{2\pi}\right)^{\frac{1}{2}Nm} e^{-\frac{1}{2}N},$$

where v_μ stands for the determinant of the v's connected with the μ-th set of x's. Thus the appropriate likelihood-ratio is given by

$$\lambda_I^{2/N} = \frac{| v^{ij} |}{v_1 \cdots v_p}.$$

It is easy to see that the value of λ_I is unaltered if we replace $x^i - a^i$ by x^i, so that we can express the probability that λ_I will lie between 0 and λ_ϵ in the form

$$I(B, \lambda_\epsilon) = \frac{B^{\frac{1}{2}N}}{\pi^{\frac{1}{2}Nm}} \int_{\lambda<\lambda_\epsilon} e^{-\sum_{\alpha=1}^{N} B_{ij} x_\alpha^i x_\alpha^j} dx_1^1 \cdots dx_N^m.$$

Furthermore, λ_I is invariant under the operation of replacing any x by a linear combination of x's belonging to the same set. And since the assumption that $\| B_{ij} \|$ is positive definite implies that the matrices $\| B_{i_\mu j_\mu} \|$ have the same property, we can transform the x's in each set among themselves by orthogonal transformations in such a way as to reduce each of the expressions

$$B_{i_\mu j_\mu} x^{i_\mu} x^{j_\mu}$$

to sums of squares. Thus we have

$$I(B, \lambda_\epsilon) = \frac{B^{*\frac{1}{2}N}}{\pi^{\frac{1}{2}Nm}} \int_{\lambda<\lambda_\epsilon} e^{-\sum_{\alpha=1}^{N} B^*_{ij} x_\alpha^i x_\alpha^j} dx_1^1 \cdots dx_N^m = I(B^*, \lambda_\epsilon), \tag{5.2}$$

where

$$B^*_{i_\mu j_\nu} = \alpha_{i_\mu}^{h_\mu} B_{h_\mu k_\nu} \alpha_{j_\nu}^{k_\nu} \qquad (h_\mu, i_\mu, j_\mu, k_\mu = m_{\mu-1} + 1, \cdots, m_\mu), \tag{5.3}$$

$$B^*_{i_\sigma i_\sigma} = 0 \qquad i_\sigma \neq j_\sigma, \tag{5.4}$$

and the subscripts on the indices indicate the sets of values over which they range; e.g., i_2 runs over the numbers corresponding to the columns of the matrix $\| B_2 \|$. From (5.3) and (5.4) it is clear that $\| B^*_{ij} \|$ reduces to a diagonal matrix when H is true.

In order to show that the test is locally unbiased, we may consider the derivatives

$$\left(\frac{\partial}{\partial B^*_{i_\mu j_\nu}} I(B^*, \lambda_\epsilon)\right)_0, \qquad \left(\frac{\partial^2}{\partial B^*_{i_\mu j_\nu} \partial B^*_{h_\sigma k_\tau}} I(B^*, \lambda_\epsilon)\right)_0, \qquad (\mu \neq \nu, \sigma \neq \tau)$$

for the B^*'s are linear functions of the B's; and the positive definiteness of one matrix of second partials implies that of the other. We have at once

$$\left(\frac{\partial B^*}{\partial B^*_{i_\mu j_\nu}}\right)_0 = 0, \qquad \left(\frac{\partial^2 B^*}{\partial B^*_{i_\mu j_\nu} \partial B_{h_\sigma k_\tau}}\right)_0 = 0, \qquad (\mu \neq \nu, \sigma \neq \tau)$$

unless the second derivative is taken twice with respect to the same B^*. Thus

$$\left(\frac{\partial I(B^*, \lambda_\epsilon)}{\partial B^*_{i_\mu j_\nu}}\right)_0 = -2 \frac{B_0^{*\frac{1}{2}N}}{\pi^{\frac{1}{2}Nm}} \int_{\lambda < \lambda_\epsilon} \sum_{\alpha=1}^{N} x_\alpha^{i_\mu} x_\alpha^{j_\nu} e^{-\sum_{\alpha=1}^{N} B^*_{ij0} x_\alpha^i x_\alpha^j} dx,$$

where the B_0^* indicates that the B's have the diagonal form associated with H. And since whenever the point $x_1^1, \cdots ; x_1^i, \cdots, x_N^i ; \cdots, x_N^m$ is in the region $\lambda \leq \lambda_\epsilon$, so also is the point $x_1^1, \cdots ; -x_1^i, \cdots, -x_N^i ; \cdots, x_N^m$ it follows that

$$\frac{\partial}{\partial B^*_{i_\mu j_\nu}} I(B_0^*, \lambda_\epsilon) = 0, \qquad (\mu \neq \nu).$$

Similar considerations show that the non-repeated second derivatives

$$\frac{\partial^2}{\partial B^*_{i_\mu j_\nu} \partial B^*_{h_\sigma k_\tau}} I(B_0^*, \lambda_\epsilon) = 4 \frac{B_0^{*\frac{1}{2}N}}{\pi^{\frac{1}{2}Nm}} \int_{\lambda < \lambda_\epsilon} \left(\sum_{\alpha=1}^{N} x_\alpha^{i_\mu} x_\alpha^{j_\nu}\right)\left(\sum_{\beta=1}^{N} x_\beta^{h_\sigma} x_\beta^{k_\tau}\right) e^{-\sum_{\alpha=1}^{N} B^*_{ij0} x_\alpha^i x_\alpha^j} dx$$

must vanish.

Finally, we must show that the repeated second derivatives are positive when evaluated at a point in ω, except of course in the trivial cases $\lambda_\epsilon = 0$, $\lambda_\epsilon = 1$, when they must be zero. In order to do this, we shall make use of the fact that the v's which go to make up λ have the Wishart distribution [17]

$$(5.5) \qquad \frac{B^{\frac{1}{2}(N-1)}}{\pi^{\frac{1}{4}(m(m-1))} \prod_{i=1}^{m} \Gamma[\frac{1}{2}(N-i)]} \cdot v^{\frac{1}{2}(N-m)-1} e^{-B_{ij}v^{ij}} dv^{11} \cdots dv^{mm}.$$

(Because of the relation $v^{ij} = v^{ji}$, only $\frac{1}{2}m(m+1)$ of the v's appear as differentials). It will be useful to have the notation

$$G(B, N-1, m) \doteq \frac{B^{\frac{1}{2}(N-1)}}{\pi^{\frac{1}{4}(m(m-1))} \prod_{i=1}^{m} \Gamma[\frac{1}{2}(N-i)]},$$

$$V(B, N-1, m) = v^{\frac{1}{2}(N-m)-1} e^{-B_{ij}v^{ij}}$$

With the aid of (5.5) we shall now compute the moments

$$E[(\lambda^{2/N})^h], \qquad h = 0, 1, \cdots,$$

for the case in which the matrix $\| B_{ij} \|$ has the form

$$\text{(5.6)} \qquad \left\| \begin{matrix} B_{11} & \cdots & B_{1m_1} & 0 & \cdots & 0B_{1m} \\ & & & & & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ B_{m_1 1} & \cdots & B_{m_1 m_1} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & & & \\ \vdots & & \vdots & & \| \bar{B} \| & \\ 0 & & & & & \\ B_{m1}0 & \cdots & 0 & & & \end{matrix} \right\|,$$

where $\| \bar{B} \|$ stands for $\| B_2 \| \dotplus \cdots \dotplus \| B_p \|$, and all other B's, except those indicated, are zero. Let us designate by $(\bar{v})$ the set of v^{ij} which correspond to the rows and columns of $\bar{B}$, and by $(v - \bar{v})$ the remaining v's. We then remark that the result of integrating (5.5) with respect to the v's in $(v - \bar{v})$ is to reduce it to the corresponding distribution for the variables in the set $\bar{v}$, thus:

$$\text{(5.7)} \qquad \begin{aligned} G(B, N - 1, m) &\int V(B, N - 1, m)\, d(v - \bar{v}) \\ &= G(\tilde{B}, N - 1, m - m_1) V(\tilde{B}, N - 1, m - m_1), \end{aligned}$$

where $\| \tilde{B}_{kl} \|$ is the inverse of the matrix obtained by inverting $\| B_{ij} \|$, and striking out the first m_1 rows and columns, that is

$$\tilde{B}^{kl} = B^{kl}, \qquad (k, l = m_1 + 1, \cdots, m).$$

Then,

$$G(B, N - 1, m) \int \frac{v^h}{v_2^h \cdots v_p^h} V(B, N - 1, m)\, d(v - \bar{v})$$

can be written as

$$\text{(5.8)} \qquad \begin{aligned} &\frac{G(B, N - 1, m)}{G(B, N - 1 + 2h, m)} \cdot G(B, N - 1 + 2h, m) \int v_2^{-h} \cdots v_p^{-h} \\ &\qquad\qquad \times V(B, N - 1 + 2h, m)\, d(v - \bar{v}) \\ &= \frac{G(B, N - 1, m)}{G(B, N - 1 + 2h, m)} G(\tilde{B}, N - 1 + 2h, m - m_1) \\ &\qquad\qquad \times v_2^{-h} \cdots v_p^{-h} V(\tilde{B}, N - 1 + 2h, m - m_1). \end{aligned}$$

It can be seen from (5.6) that

$$\| \tilde{B} \| = \| B_2 \| + \cdots \dotplus \| B_{p-1} \| \dotplus \| \tilde{B}_p \|$$

since of all the rows and columns of $\| B_{ij} \|$ which are involved in $\| \bar{B} \|$ it is only the last in which a non zero element appears outside of the blocks $\| B_2 \|$, $\cdots$, $\| B_p \|$. Consequently, the v's corresponding to the determinants $v_2, \cdots,$

v_p are independently distributed, so that if in (5.8) we integrate out all the remaining v's but these, we shall be left with a product of factors

$$\frac{G(B, N-1, m)}{G(B, N-1+2h, m)} \cdot \prod_{t=2}^{p-1} \frac{G(B_t, N-1+2h, k_t)}{G(B_t, N-1, k_t)}$$
$$\times G(B_t, N-1, k_t) v_t^{-h} V(B_t, N-1+2h, k_t)$$
$$\times \frac{G(\tilde{B}_p, N-1+2h, k_p)}{G(\tilde{B}_p, N-1, k_p)} \cdot G(\tilde{B}_p, N-1, k_p) v_p^{-h} V(\tilde{B}_p, N-1+2h, k_p),$$

where k_μ stands for the order of $\| B_\mu \|$. And this, when integrated with respect to the v's in $v_2, \cdots, v_p$, yields

$$\frac{G(B, N-1, m)}{G(B, N-1+2h, m)} \cdot \prod_{t=2}^{p-1} \frac{G(B_t, N-1+2h, k_t)}{G(B_t, N-1, k_t)} \times \frac{G(\tilde{B}_p, N-1+2h, k_p)}{G(\tilde{B}_p, N-1, k_p)},$$

which, because of the definition of the G's, reduces to

$$\prod_{i=1}^{m} \frac{\Gamma[\frac{1}{2}(N-i)+h]}{\Gamma[\frac{1}{2}(N-i)]} \cdot \prod_{t=2}^{p} \prod_{i=1}^{k_t} \frac{\Gamma[\frac{1}{2}(N-i)]}{\Gamma[\frac{1}{2}(N-i)+h]} \times B^{-h} B_2^h \cdots B_{p-1}^h \tilde{B}_p^h.$$

Denoting the product of ratios of Γ's by K_h, and recalling the form of $\| B_{ij} \|$, we therefore have

$$E\left[\frac{v^h}{v_2^h \cdots v_p^h}\right] = K_h \tilde{B}_p^h B'^{-h} \tag{5.9}$$

with

$$\| B' \| = \left\| \begin{array}{cccccc} B_{11} & \cdots & B_{1m_1} & 0 & \cdots & 0 B_{1m} \\ & & & & & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ B_{m_1 1} & \cdots & B_{m_1 m_1} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & & & \\ \vdots & & \vdots & & \| B_p \| & \\ 0 & & & & & \\ B_{m1} 0 & \cdots & 0 & & & \end{array} \right\|$$

But it is not difficult to see that under the condition (5.6), the matrix $\| \tilde{B}_p \|$ is also the inverse of the matrix obtained by striking out the first m rows and columns in the inverse of $\| B' \|$. Making use of this relation, we can apply the Jacobi theorem to (5.9), and put that expression in the form

$$E\left[\frac{v^h}{v_2^h \cdots v_p^h}\right] = K_h B_1^{-h},$$

where $\| B_1 \|$ is the matrix in the upper left hand corner of $\| B' \|$, namely $\| B_{i_1 j_1} \|$.

Let the subscript β on a B stand for the result of replacing $B_{i_1 j_1}$ by $B_{i_1 j_1} + \beta_{i_1 j_1}$. For sufficiently small values of the β's the matrix $\| B_{ij\beta} \|$ will still be positive definite, so that we shall have

$$\frac{B_\beta^{\frac{1}{2}(N-1)}}{\pi^{\frac{1}{4}(m(m-1))} \prod_{i=1}^{m} \Gamma[\frac{1}{2}(N-i)]} \int \frac{v^h}{v_2^h \cdots v_p^h} v^{\frac{1}{2}(N-m)-1} e^{-B_{ij\beta}v^{ij}} dv = K_h B_{1\beta}^{-h},$$

which we can put in the form

$$K' \int \frac{v^h}{v_2^h \cdots v_p^h} v^{\frac{1}{2}(N-m)-1} e^{-B_{ij\beta}v^{ij}} dv = \frac{K_h}{B_{1\beta}^h B_\beta^{\frac{1}{2}(N-1)}}. \tag{5.10}$$

Wilks [13] has shown how to generate moments of determinants by the device of replacing $\beta_{i_1 j_1}$ by $\beta_{i_1 j_1} + \xi_{i_1}\xi_{j_1}$, and integrating with respect to the ξ's from $-\infty$ to ∞. Applying this process $2h$ times to the left hand side of (5.10) gives

$$\pi^{t_1 h} K' \int \left(\frac{v}{v_1 \cdots v_p}\right)^h V(B_\beta, N-1, m)\, dv,$$

which when multiplied by $\pi^{-t_1 h} B^{\frac{1}{2}(N-1)}$ yields

$$E[(\lambda^{2/N})^h]$$

when the β's are set equal to zero.

To obtain the value of this expression, we may perform the same operations on the right hand side of (5.10). But before so doing, we shall put B_β in a more convenient form. We have

$$B_\beta = B_{1\beta} \cdot \bar{B} - B_{1m}^2 \cdot \bar{B}\bar{B}^{mm} \cdot B_{1\beta}'^{11},$$

where $\bar{B}^{mm}$ is the inverse element of B_{mm} in $\| \bar{B} \|$, and $B_{1\beta}'^{11}$ is the cofactor of $B_{11\beta}$ in $B_{1\beta}$, the result being obtained by expanding B_β according to minors of the first row and first column. Similarly,

$$B = B_1 \cdot \bar{B} - B_{1m}^2 \cdot \bar{B}\bar{B}^{mm} \cdot B_1'^{11}. \tag{5.11}$$

From (5.11) we have

$$\frac{B}{B_1 \cdots B_p} = 1 - B_{1m}^2 \bar{B}^{mm} \cdot \frac{B_1'^{11}}{B_1},$$

so that if we put $B \cdot B_1^{-1} \cdots B_p^{-1} = \Lambda$, we find that

$$B_\beta = B_{1\beta} \cdot \bar{B} \left\{ 1 - B_{1m}^2 \bar{B}^{mm} \cdot \frac{B_1'^{11}}{B_1} \cdot \frac{B_1}{B_1'^{11}} \cdot \frac{B_{1\beta}'^{11}}{B_{1\beta}} \right\}$$

$$= B_{1\beta}\bar{B} \left\{ 1 - \frac{B_1}{B_1'^{11}} (1 - \Lambda) \frac{B_{1\beta}'^{11}}{B_{1\beta}} \right\}.$$

Thus the result of multiplying (5.10) through by $B^{\frac{1}{2}(N-1)}$ (where no β's are substituted in this determinant) can be put in the form

$$\left(\frac{B}{B_2 \cdots B_p}\right)^{\frac{1}{2}(N-1)} \left\{1 - \frac{B_1}{B_1'^{11}}(1-\Lambda)\frac{B_{1\beta}'^{11}}{B_{1\beta}}\right\}^{-\frac{1}{2}(N-1)} B_{1\beta}^{-h}. \tag{5.12}$$

Expanding the expression in curled brackets, we get

$$\Lambda^{\frac{1}{2}(N-1)} B_1^{\frac{1}{2}(N-1)} \sum_{\nu=0}^{\infty} \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\nu!\,\Gamma[\frac{1}{2}(N-1)]} B_1^{\nu} \left(\frac{B_{1\beta}'^{11}}{B_1'^{11}}\right)^{\nu} B_{1\beta}^{-[\frac{1}{2}(N-1)+h+\nu]} (1-\Lambda)^{\nu}.$$

If we let $B_{1\beta t}$ stand for the result of replacing B_{11} by $B_{11} - t$ in $B_{1\beta}$, we can write this as

$$\begin{aligned} &\Lambda^{\frac{1}{2}(N-1)} \sum_{\nu=0}^{\infty} \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\nu!\,\Gamma[\frac{1}{2}(N-1)]} (1-\Lambda)^{\nu} (B_1'^{11})^{-\nu} B_1^{\frac{1}{2}(N-1)+\nu} \\ &\qquad \times \frac{\Gamma[\frac{1}{2}(N-1)+h]}{\Gamma[\frac{1}{2}(N-1)+h+\nu]} \frac{\partial^{\nu}}{\partial t^{\nu}} B_{1\beta t}^{-[\frac{1}{2}(N-1)+h]} \end{aligned} \tag{5.13}$$

the derivatives being evaluated at $t = 0$.

Now Wilks' results show that the operation of introducing $\beta_{i_1 j_1} + \xi_{i_1}\xi_{j_1}$ into $B_{1\beta t}$ to replace $\beta_{i_1 j_1}$ and integrating with respect to the ξ's, when repeated $2h$ times on $B_{1\beta t}^{-[\frac{1}{2}(N-1)+h]}$, produces

$$\pi^{m_1 h} B_{1t}^{-\frac{1}{2}(N-1)} \prod_{i=1}^{m_1} \frac{\Gamma[\frac{1}{2}(N-i)]}{\Gamma[\frac{1}{2}(N-i)+h]}$$

when the β's are finally set equal to zero. Reversing the order of summation, differentiation and integration in (5.13), we thus obtain

$$\begin{aligned} &\pi^{m_1 h} \prod_{i=1}^{m_1} \frac{\Gamma[\frac{1}{2}(N-i)]}{\Gamma[\frac{1}{2}(N-i)+h]} \Lambda^{\frac{1}{2}(N-1)} \sum_{\nu=0}^{\infty} \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\nu!\,\Gamma[\frac{1}{2}(N-1)]} \\ &\qquad \times (1-\Lambda)^{\nu} (B_1'^{11})^{-\nu} B_1^{\frac{1}{2}(N-1)+\nu} \frac{\Gamma[\frac{1}{2}(N-1)+h]}{\Gamma[\frac{1}{2}(N-1)+h+\nu]} \left(\frac{\partial^{\nu}}{\partial t^{\nu}} B_{1t}^{-\frac{1}{2}(N-1)}\right)_0. \end{aligned} \tag{5.14}$$

Now

$$\left(\frac{\partial^{\nu}}{\partial t^{\nu}} B_{1t}^{-\frac{1}{2}(N-1)}\right)_0 = \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\Gamma[\frac{1}{2}(N-1)]} \cdot (B_1'^{11})^{\nu} B_1^{-[\frac{1}{2}(N-1)+\nu]},$$

so that (5.14) becomes

$$\begin{aligned} &\pi^{m_1 h} \prod_{i=1}^{m_1} \frac{\Gamma[\frac{1}{2}(N-i)]}{\Gamma[\frac{1}{2}(N-i)+h]} \Lambda^{\frac{1}{2}(N-1)} \sum_{\nu=0}^{\infty} \frac{\Gamma[\frac{1}{2}(N-1)]}{\nu!\,\Gamma[\frac{1}{2}(N-1)]} \\ &\qquad \times (1-\Lambda)^{\nu} \frac{\Gamma[\frac{1}{2}(N-1)+h]}{\Gamma[\frac{1}{2}(N-1)+h+\nu]} \cdot \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\Gamma[\frac{1}{2}(N-1)]}. \end{aligned}$$

From this it appears that the h-th moment of $\lambda_I^{2/N}$ is given by

$$E[(\lambda^{2/N})^h] = \prod_{i=1}^{m} \frac{\Gamma[\frac{1}{2}(N-i)+h]}{\Gamma[\frac{1}{2}(N-i)]} \cdot \prod_{t=1}^{p} \prod_{i=1}^{k_t} \frac{\Gamma[\frac{1}{2}(N-i)]}{\Gamma[\frac{1}{2}(N-i)+h]}$$

$$(5.15) \qquad \times \Lambda^{\frac{1}{2}(N-1)} \sum_{\nu=0}^{\infty} (1-\Lambda)^\nu \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\nu!\,\Gamma[\frac{1}{2}(N-1)]}$$

$$\times \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\Gamma[\frac{1}{2}(N-1)+h+\nu]} \cdot \frac{\Gamma[\frac{1}{2}(N-1)+h]}{\Gamma[\frac{1}{2}(N-1)]}.$$

A considerable amount of cancellation will take place in (5.15), for m is greater than any k_t. Suppose the largest k_t is $k_{t'}$. Then we can cancel its product into the first one, with the assurance that there will be at least one factor

$$(5.16) \qquad \frac{\Gamma[\frac{1}{2}(N-1)]}{\Gamma[\frac{1}{2}(N-1)+h]}$$

to cancel the corresponding factor under the summation sign. Hence we have

$$E[(\lambda^{2/N})^h] = \prod_{i=k_{t'}+1}^{m} \frac{\Gamma[\frac{1}{2}(N-i)+h]}{\Gamma[\frac{1}{2}(N-i)]} \cdot {\prod_{t=1}^{p}}' \, {\prod_{i=1}^{k_t}}'' \frac{\Gamma[\frac{1}{2}(N-i)]}{\Gamma[\frac{1}{2}(N-i)+h]}$$

$$(5.17) \qquad \times \Lambda^{\frac{1}{2}(N-1)} \sum_{\nu=0}^{\infty} (1-\Lambda)^\nu \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\nu!\,\Gamma[\frac{1}{2}(N-1)]} \cdot \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\Gamma[\frac{1}{2}(N-1)+h+\nu]},$$

where Π' indicates that t' has been omitted, and Π'' indicates that one factor (5.16) has been cancelled. Then we can take out the factor $i = m$ in the first product, putting it under the summation sign, where, together with the final factor in each term of the sum, it gives rise to the combination

$$\frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\Gamma[\frac{1}{2}(N-m)]\Gamma[\frac{1}{2}(m-1)+\nu]} \cdot \frac{\Gamma[\frac{1}{2}(N-m)+h]\Gamma[\frac{1}{2}(m-1)+\nu]}{\Gamma[\frac{1}{2}(N-1)+h+\nu]}.$$

After making this reduction, we obtain

$$E[(\lambda^{2/N})^h] = \prod_{i=k_{t'}+1}^{m-1} \frac{\Gamma[\frac{1}{2}(N-i)+h]}{\Gamma[\frac{1}{2}(N-i)]} \cdot {\prod_{t=2}^{p}}' \, {\prod_{i=1}^{k_t}}' \frac{\Gamma[\frac{1}{2}(N-i)]}{\Gamma[\frac{1}{2}(N-i)+h]}$$

$$(5.18) \qquad \times \Lambda^{\frac{1}{2}(N-1)} \sum_{\nu=0}^{\infty} (1-\Lambda)^\nu \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\nu!\,\Gamma[\frac{1}{2}(N-1)]} \frac{B[\frac{1}{2}(N-m)+h, \frac{1}{2}(m-1)+\nu]}{B[\frac{1}{2}(N-m), \frac{1}{2}(m-1)+\nu]}.$$

The products of ratios in the first part of (5.18) are of the type discussed by Wilks in connection with integral equations of type B [12]. It follows from his results that $\lambda_I^{2/N}$ is distributed like the product

$$z \cdot \theta_1 \cdots \theta_{m'} \qquad (m' = m - k_{t'} - 1),$$

where z and the θ's are independently distributed, with the distribution of the θ's given by

$$f(\theta_1, \cdots, \theta_{m'}) = \prod_{i=1}^{m'} \frac{\Gamma(c_i)}{\Gamma(b_i)\Gamma(c_i - b_i)} \cdot \theta_i^{b_i - 1}(1-\theta_i)^{c_i - b_i - 1},$$

where the b_i and c_i are constants which depend on N, m, and the sizes of the blocks, but not on Λ, and the distribution of z is given by

$$F(z) = \Lambda^{\frac{1}{2}(N-1)} \sum_{\nu=0}^{\infty} (1-\Lambda)^{\nu} \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\nu!\,\Gamma[\frac{1}{2}(N-1)]} \cdot \frac{z^{\frac{1}{2}(N-m)-1}(1-z)^{\frac{1}{2}(m-1)+\nu-1}}{B[\frac{1}{2}(N-m), \frac{1}{2}(m-1)+\nu]}.$$

Consequently, the probability that λ lies between zero and λ_ϵ is

$$J(\Lambda, \lambda_\epsilon) = \Lambda^{\frac{1}{2}(N-1)} \int_S \sum_{\nu=0}^{\infty} (1-\Lambda)^{\nu} \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\nu!\,\Gamma[\frac{1}{2}(N-1)]} \times f(\theta) \frac{z^{\frac{1}{2}(N-m)-1}(1-z)^{\frac{1}{2}(m-1)+\nu-1}}{B[\frac{1}{2}(N-m), \frac{1}{2}(m-1)+\nu]}\, dz\, d\theta,$$

where the integral is to be extended over the region

$$S:\ 0 \leq z\cdot\theta_1 \cdots \theta_{m'} < \lambda_\epsilon^{2/N}, \qquad 0 \leq \theta_i \leq 1, \qquad 0 \leq z \leq 1.$$

Let us integrate first with respect to z and then with respect to the θ's; we have

$$J(\Lambda, \lambda_\epsilon) = \int_{S_\theta} \Lambda^{\frac{1}{2}(N-1)} \sum_{\nu=0}^{\infty} (1-\Lambda)^{\nu} \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\nu!\,\Gamma[\frac{1}{2}(N-1)]} \times \frac{B'[\frac{1}{2}(N-m), \frac{1}{2}(m-1)+\nu; \varphi]}{B[\frac{1}{2}(N-m), \frac{1}{2}(m-1)+\nu]} f(\theta)\, d\theta, \tag{5.19}$$

where S_θ is the set $\Pi\theta_i < \lambda_\epsilon^{2/N}$, $0 \leq \theta_i \leq 1$, and

$$\begin{aligned} B'(u, v, \varphi) &= \int_0^{\varphi} z^{u-1}(1-z)^{v-1}\, dz \\ &= \int_{1-\varphi}^{1} z^{v-1}(1-z)^{u-1}\, dz = B(v, u, 1-\varphi), \end{aligned} \tag{5.20}$$

$\varphi(\theta)$ being the upper limit for z for fixed θ. It is clear that the subset of s_θ for which $\varphi(\theta) < 1$ will not be of measure zero in the θ-space, since we assume that $0 < \lambda_\epsilon < 1$.

The relation between (5.19) and the corresponding expression for the multiple correlation coefficient without fixed variates—the case $\bar{y} = 0$ in (4.4)—may be clearer if we put

$$\bar{\rho} = 1 - \Lambda = B_{1m}^2 \bar{B}^{mm} B_1^{11}, \tag{5.21}$$

where $\bar{B}^{mm}$ is the inverse of B_{mm} in $\| \bar{B} \|$, and B_1^{11} is the inverse of B_{11} in $\| B_1 \|$. Then the required probability of rejection when $\bar{\rho}$ has any fixed value is

$$I(\bar{\rho}, 1-\lambda_\epsilon^{2/N}) = \int_{S_\theta} \sum_{\nu=0}^{\infty} \frac{\bar{\rho}^{\nu}}{\nu!} (1-\bar{\rho})^{\frac{1}{2}(N-1)} \frac{\Gamma[\frac{1}{2}(N-1)+\nu]}{\Gamma[\frac{1}{2}(N-1)]} \times \frac{B[\frac{1}{2}(m-1)+\nu, \frac{1}{2}(N-m), 1-\varphi]}{B[\frac{1}{2}(m-1)+\nu, \frac{1}{2}(N-m]} f(\theta)\, d\theta,$$

where we have used the relation (5.20) between the incomplete Beta functions. Differentiating with respect to $\bar{\rho}$ before performing the integration with respect

to the θ's, we find by a computation similar to that in section 4 that each term in the series is positive except where $\varphi(\theta) = 1$; so that we have

$$\frac{\partial I}{\partial \bar{\rho}}(\bar{\rho}, 1 - \lambda_\epsilon^{2/N}) > 0 \qquad (\lambda_\epsilon \neq 1, 0).$$

And by (5.21), we then have

$$\frac{\partial^2 I}{\partial B_{1m}^2} > 0.$$

Since the argument is clearly independent of which $B_{i_\mu j_\nu}$ $(\mu \neq \nu)$ we take, it follows that the test is locally unbiased. We have therefore proved:

THEOREM III. *If $x^1, \cdots, x^m$ have the joint normal distribution* (5.1), *then the likelihood ratio test for the hypothesis that the x's are independent in sets is locally unbiased.*

In certain types of statistical material it may be important to consider, not the independence of the x's themselves, but of their deviations from regression functions. For example, in the case of several related time series, it may be desirable to eliminate the trend of each x^i by means of, say, a second degree polynomial in t. Consider then in general a population whose distribution function is of the form

$$\frac{B^{\frac{1}{2}}}{\pi^{\frac{1}{2}m}} e^{-B_{ij}(x^i - C^i_\mu x^\mu)(x^j - C^j_\nu x^\nu)} \qquad (\mu, \nu = m + 1, \cdots, m + q)$$

with unknown B_{ij} and C^i_μ. The likelihood ratio for testing the hypothesis H_I that the sets of deviations

$$x^1 - C^1_\mu x^\mu, \cdots, x^{m_1} - C^{m_1}_\mu x^\mu; \cdots; x^{m_{p-1}+1} - C^{m_{p-1}+1}_\mu x^\mu, \cdots, x^m - C^m_\mu x^\mu$$

are independent is

$$\lambda_I = \left\{\frac{|d^{ij}|}{d_1 \cdots d_p}\right\}^{\frac{1}{2}N}$$

where

$$d^{ij} = \Sigma(x^i_\alpha - \hat{C}^i_\mu x^\mu_\alpha)(x^j_\alpha - \hat{C}^j_\nu x^\nu_\alpha)$$

and $\hat{C}^i_\mu$ is the usual least squares estimate of C^i_μ, given by

$$\hat{C}^i_\mu a^{\mu\nu} = a^{i\nu}$$

with

$$a^{rs} = \Sigma x^r_\alpha x^s_\alpha \qquad (r, s = 1, \cdots, m + q).$$

An examination of the characteristic function of the d^{ij} shows that their distribution law is the same as that of the v^{ij} of the preceding discussion, except for the fact that $N - 1$ is replaced by $N - q$. Consequently the above results on freedom from bias, and also those of the next section, apply equally well to the λ_I test for the independence of deviations from regression functions.

6. **On the moments of $\lambda_I^{2/N}$.** Although we have succeeded in proving the unbiased nature of the preceding test only in the local sense, we can show that the moments of the criterion $\lambda_I^{2/N}$ have a property which seems very closely related to that of furnishing a completely unbiased test. For it can be shown that each of the quantities

$$E[(\lambda^{2/N})^h] \qquad h = \tfrac{1}{2}, 1, 1\tfrac{1}{2}, \cdots$$

is greater, when H_I is true than when any alternative H' holds. It will perhaps be sufficient to prove this statement in detail for the case where $h = 1$ and where H_I is the hypothesis that the matrix $\| B_{ij} \|$ has the form $\| \breve{B}_0 \| \dotplus \|B_{i_3 j_3}\|$:

$$\left\| \begin{matrix} \begin{matrix} B_{11} B_{12} \\ B_{21} B_{22} \end{matrix} & 0 & 0 \\ 0 & \begin{matrix} B_{33} B_{34} \\ B_{43} B_{44} \end{matrix} & 0 \\ 0 & 0 & \| B_{i_3 j_3} \| \end{matrix} \right\| .$$

in the notation of the preceding section we then have

$$i_1, j_1 = 1, 2,; \qquad i_2, j_2 = 3, 4; \qquad i_3, j_3 = 5, \cdots, m.$$

Even when H is not true we find that

$$(6.1) \quad E[| v^{ij} |^h | v^{i_3 j_3} |^{-h}] = \frac{G(B, N-1, m)}{G(B, N-1+2h, m)} \cdot \frac{G(\tilde{B}, N-1+2h, m-4)}{G(\tilde{B}, N-1, m-4)},$$

where $\tilde{B}^{i_3 j_3} = B^{i_3 j_3}$. Using the definition of the G's in section 5 and the Jacobi theorem, we can write (6.1) in the form

$$E[| v^{ij} |^h | v^{i_3 j_3} |^{-h}] = K_h \breve{B}^{-h}$$

where $\breve{B}$ is the determinant of the matrix composed of the first four rows and columns of $\| B_{ij} \|$. In the general case we therefore have

$$\| \breve{B} \| = \left\| \begin{matrix} B_{11} & B_{12} & B_{13} & B_{14} \\ B_{21} & B_{22} & B_{23} & B_{24} \\ B_{31} & B_{32} & B_{33} & B_{34} \\ B_{41} & B_{42} & B_{43} & B_{44} \end{matrix} \right\| .$$

Thus if we set $h = 1$, and replace $B_{i_1 j_1}$ and $B_{i_2 j_2}$ by $B_{i_1 j_1} + \xi_{i_1}^{(1)} \xi_{j_1}^{(1)} + \xi_{i_1}^{(2)} \xi_{j_1}^{(2)}$ and $B_{i_2 j_2} + \xi_{i_2}^{(3)} \xi_{j_2}^{(3)} + \xi_{i_2}^{(4)} \xi_{j_2}^{(4)}$ respectively, indicating this replacement by a prime, we obtain

$$(6.2) \qquad E[(\lambda^{2/N})^1] = K_1 \int B^{\frac{1}{2}(N-1)} B'^{-\frac{1}{2}(N-1)} \breve{B}'^{-1} d\xi.$$

Treating B' as a bordered determinant, we can reduce it to

$$\begin{aligned} B' &= B_{(123)}(1 + B^{i_2 j_2}_{(123)}\xi^{(4)}_{i_2}\xi^{(4)}_{j_2}) \\ &= B_{(12)}(1 + B^{i_2 j_2}_{(12)}\xi^{(3)}_{i_2}\xi^{(3)}_{j_2})(1 + B^{i_2 j_2}_{(123)}\xi^{(4)}_{i_2}\xi^{(4)}_{j_2}) \\ &= B_{(1)}(1 + B^{i_1 j_1}_{(1)}\xi^{(2)}_{i_1}\xi^{(2)}_{j_1})(1 + B^{i_2 j_2}_{(12)}\xi^{(3)}_{i_2}\xi^{(3)}_{j_2})(1 + B^{i_2 j_2}_{(123)}\xi^{(4)}_{i_2}\xi^{(4)}_{j_2}) \\ &= B(1 + B^{i_1 j_1}\xi^{(1)}_{i_1}\xi^{(1)}_{j_1})(1 + B^{i_1 j_1}_{(1)}\xi^{(2)}_{i_1}\xi^{(2)}_{j_1})(1 + B^{i_2 j_2}_{(12)}\xi^{(3)}_{i_2}\xi^{(3)}_{j_2})(1 + B^{i_2 j_2}_{(123)}\xi^{(4)}_{i_2}\xi^{(4)}_{j_2}), \end{aligned}$$

where the subscripts on the B's indicate the sets of ξ's still contained in the determinants, and $\| B^{ij} \| = \| B_{ij} \|^{-1}$. Similarly,

$$(6.4)\quad \begin{aligned} \breve{B}' &\\ &= \breve{B}(1 + \breve{B}^{i_1 j_1}\xi^{(1)}_{i_1}\xi^{(1)}_{j_1})(1 + \breve{B}^{i_1 j_1}_{(1)}\xi^{(2)}_{i_1}\xi^{(2)}_{j_1})(1 + \breve{B}^{i_2 j_2}_{(12)}\xi^{(3)}_{i_2}\xi^{(3)}_{j_2})(1 + \breve{B}^{i_2 j_2}_{(123)}\xi^{(4)}_{i_2}\xi^{(4)}_{j_2}), \end{aligned}$$

the inverse now being taken with respect to $\| \breve{B} \|$.

But between, say, $\breve{B}^{i_2 j_2}_{(12)}$ and $B^{i_2 j_2}_{(12)}$, there is the relation

$$(6.5)\qquad \breve{B}^{i_2 j_2}_{(12)} = B^{i_2 j_2}_{(12)} - B^{i_2 i_3}_{(12)} B_{(12)\, i_3 j_3} B^{j_3 j_2}_{(12)},$$

where $\| B_{(12)\, i_3 j_3} \| = \| B^{i_3 j_3}_{(12)} \|^{-1}$, that is, the inverse of the matrix obtained by deleting the first four rows and columns of $\| B^{ij}_{(12)} \|$. Consequently

$$\breve{B}^{i_2 j_2}_{(12)}\xi^{(3)}_{i_2}\xi^{(3)}_{j_2} \leq B^{i_2 j_2}_{(12)}\xi^{(3)}_{i_2}\xi^{(3)}_{j_2}$$

with equality holding only for those values of the ξ's for which

$$\xi^{(3)}_{i_2} B^{i_2 i_3}_{(12)} = 0 \qquad i_3 = 5, \cdots, m.$$

And this set of ξ's will not make up the entire ξ space unless $\| B_{ij} \| = \| \breve{B} \| \dotplus \| B_{i_3 j_3} \|$. Applying the same kind of reasoning to the other quadratic forms in (6.4), we can therefore show that

$$\begin{aligned} &\int B^{\frac{1}{2}(N-1)} B'^{-\frac{1}{2}(N-1)} \breve{B}'^{-1}\, d\xi \\ &\qquad \leq \breve{B}^{-1} \int (1 + \breve{B}^{i_1 j_1}\xi^{(1)}_{i_1}\xi^{(1)}_{j_1})^{-\frac{1}{2}(N+1)} \cdots (1 + \breve{B}^{i_2 j_2}_{(123)}\xi^{(4)}_{i_2}\xi^{(4)}_{j_2})^{-\frac{1}{2}(N+1)}\, d\xi. \end{aligned}$$

The last form can be reduced to a sum of squares with unit coefficients by a linear transformation of the $\xi^{(4)}$'s; thus

$$(6.6)\quad \begin{aligned} &\int B^{\frac{1}{2}(N-1)} B'^{-\frac{1}{2}(N-1)} \breve{B}'^{-1}\, d\xi \\ &\qquad \leq \breve{B}^{-1} \int | \breve{B}^{i_2 j_2}_{(123)} |^{-\frac{1}{2}} (1 + \breve{B}^{i_1 j_1}\xi^{(1)}_{i_1}\xi^{(1)}_{j_1})^{-\frac{1}{2}(N-1)} \cdots (1 + \Sigma\xi^{(4)}_{i_2}\xi^{(4)}_{i_2})^{-\frac{1}{2}(N+1)}\, d\xi. \end{aligned}$$

And by making use of the fact that

$$\breve{B}^{i_2 j_2}_{(123)} = \breve{B}^{-1}_{(123)} \cdot | B_{(12)\, i_2 j_2} |,$$

we can express the right-hand side of (6.6) as

$$\breve{B}^{-1}\int \breve{B}^{\frac{1}{2}}_{(123)}\cdot|B_{(12)i_1j_1}|^{-\frac{1}{2}}(1+\breve{B}^{i_1j_1}\xi^{(1)}_{i_1}\xi^{(1)}_{j_1})^{-\frac{1}{2}(N+1)}\cdots(1+\Sigma\xi^{(4)}_{i_2}\xi^{(4)}_{i_2})^{-\frac{1}{2}(N+1)}\,d\xi.$$

This in turn becomes [c. f. (6.4)]

$$\begin{aligned}\breve{B}^{-\frac{1}{2}}\int|B_{(12)i_1j_1}|^{-\frac{1}{2}}&(1+\breve{B}^{i_1j_1}\xi^{(1)}_{i_1}\xi^{(1)}_{j_1})^{-\frac{1}{2}N}(1+\breve{B}^{i_1j_1}_{(1)}\xi^{(2)}_{i_1}\xi^{(2)}_{j_1})^{-\frac{1}{2}N}\\ &\times(1+\breve{B}^{i_2j_2}_{(12)}\xi^{(3)}_{i_2}\xi^{(3)}_{j_2})^{-\frac{1}{2}N}(1+\Sigma\xi^{(4)}_{i_2}\xi^{(4)}_{j_2})^{-\frac{1}{2}(N+1)}\,d\xi\\ &=\int|B_{(12)i_1j_1}|^{-1}(1+\breve{B}^{i_1j_1}\xi^{(1)}_{i_1}\xi^{(1)}_{j_1})^{-\frac{1}{2}N}(1+\breve{B}^{i_1j_1}_{(1)}\xi^{(2)}_{i_1}\xi^{(2)}_{j_1})^{-\frac{1}{2}(N-1)}\\ &\times(1+\Sigma\xi^{(3)}_{i_2}\xi^{(3)}_{i_2})^{-\frac{1}{2}N}(1+\Sigma\xi^{(4)}_{i_2}\xi^{(4)}_{i_2})^{-\frac{1}{2}(N+1)}\,d\xi.\end{aligned}$$

At this stage we can write

$$|B_{(12)i_1j_1}|=|B_{i_1j_1}|(1+B^{*i_1j_1}\xi^{(1)}_{i_1}\xi^{(1)}_{j_1})(1+B^{*i_1j_1}_{(1)}\xi^{(2)}_{i_1}\xi^{(2)}_{j_1}),$$

where $\|B^{*i_1j_1}_{(1)}\|=\|B_{(1)i_1j_1}\|^{-1}$, and apply the relation

$$B^{*i_1j_1}_{(1)}=\breve{B}^{i_1j_1}_{(1)}-\breve{B}^{i_1i_2}_{(1)}\breve{B}_{(1)i_2j_2}\breve{B}^{j_2j_1}_{(1)},\qquad\|\breve{B}_{(1)i_2j_2}\|=\|\breve{B}^{i_2j_2}_{(1)}\|^{-1}.$$

Therefore,

$$B^{*i_1j_1}_{(1)}\xi^{(2)}_{i_1}\xi^{(2)}_{j_1}<\breve{B}^{i_1j_1}_{(1)}\xi^{(2)}_{i_1}\xi^{(2)}_{j_1},$$

unless $\xi^{(2)}_{i_1}\breve{B}^{i_1i_2}_{(1)}=0$ $(i_2=3,4)$. We can thus continue as follows

$$\begin{aligned}&\int B^{\frac{1}{2}(N-1)}B'^{-\frac{1}{2}(N-1)}\breve{B}'^{-1}\,d\xi\\ &\quad\le|B_{i_1j_1}|^{-1}\int(1+B^{*i_1j_1}\xi^{(1)}_{i_1}\xi^{(2)}_{j_1})^{-\frac{1}{2}(N+1)}(1+B^{*i_1j_1}_{(1)}\xi^{(2)}_{i_1}\xi^{(2)}_{j_1})^{-\frac{1}{2}(N+1)}\\ &\qquad\times(1+\Sigma\xi^{(3)}_{i_2}\xi^{(3)}_{i_2})^{-\frac{1}{2}N}(1+\Sigma\xi^{(4)}_{i_2}\xi^{(4)}_{i_2})^{-\frac{1}{2}(N+1)}\,d\xi.\end{aligned}$$

Transforming the $\xi^{(2)}$'s, we get

$$\begin{aligned}|B_{i_1j_1}|^{-1}\int|B^{*i_1j_1}_{(1)}|^{-\frac{1}{2}}&(1+B^{*i_1j_1}\xi^{(1)}_{i_1}\xi^{(1)}_{j_1})^{-\frac{1}{2}(N+1)}(1+\Sigma\xi^{(2)}_{i_1}\xi^{(2)}_{i_1})^{-\frac{1}{2}(N+1)}\\ &\times(1+\Sigma\xi^{(3)}_{i_2}\xi^{(3)}_{i_2})^{-\frac{1}{2}N}(1+\Sigma\xi^{(4)}_{i_2}\xi^{(4)}_{i_2})^{-\frac{1}{2}(N+1)}\,d\xi.\end{aligned}$$

Since $|B^{*i_1j_1}_{(1)}|^{-1}=|B_{(1)i_1j_1}|$, this becomes

$$\begin{aligned}|B_{i_1j_1}|^{-\frac{1}{2}}\int&(1+B^{*i_1j_1}\xi^{(1)}_{i_1}\xi^{(1)}_{j_1})^{-\frac{1}{2}N}(1+\Sigma\xi^{(2)}_{i_1}\xi^{(2)}_{i_1})^{-\frac{1}{2}(N+1)}\\ &\times(1+\Sigma\xi^{(3)}_{i_2}\xi^{(3)}_{i_2})^{-\frac{1}{2}N}(1+\Sigma\xi^{(4)}_{i_2}\xi^{(4)}_{j_2})^{-\frac{1}{2}(N+1)}\,d\xi\\ &=\int(1+\Sigma\xi^{(1)}_{i_1}\xi^{(1)}_{i_1})^{-\frac{1}{2}N}(1+\Sigma\xi^{(2)}_{i_1}\xi^{(2)}_{i_1})^{-\frac{1}{2}(N+1)}\\ &\times(1+\Sigma\xi^{(3)}_{i_2}\xi^{(3)}_{i_2})^{-\frac{1}{2}N}(1+\Sigma\xi^{(4)}_{i_2}\xi^{(4)}_{i_2})^{-\frac{1}{2}(N+1)}\,d\xi.\end{aligned}$$

Collecting these results, we finally obtain

$$
\begin{aligned}
& K_1 \int B^{\frac{1}{2}(N-1)} B'^{-\frac{1}{2}(N-1)} \breve{B}'^{-1} \, d\xi \\
(6.7) \qquad & \leq K_1 \int (1 + \Sigma \xi_{i_1}^{(1)} \xi_{i_1}^{(1)})^{-\frac{1}{2}N} (1 + \Sigma \, \xi_{i_1}^{(2)} \xi_{i_1}^{(2)})^{-\frac{1}{2}(N+1)} \\
& \qquad \times (1 + \Sigma \xi_{i_2}^{(3)} \xi_{i_2}^{(3)})^{-\frac{1}{2}N} (1 + \Sigma \xi_{i_2}^{(4)} \xi_{i_2}^{(4)})^{-\frac{1}{2}(N+1)} \, d\xi
\end{aligned}
$$

with equality only in case H_I is true. But the right side of (6.7) is the first moment of $\lambda_I^{2/N}$ computed under the hypothesis H_I, while the left side gives the corresponding moment in the general case.

The possibility of carrying out this reduction for the case in which the matrix $\| \breve{B} \|$ has more than two blocks, or blocks of unequal size, seems sufficiently clear. And to obtain higher moments, we have only to introduce the proper number of ξ's into each set. We then have:

THEOREM IIIa. *Let λ_i be the likelihood ratio appropriate to testing the hypothesis H_I that the normally distributed variates $x^1, \cdots, x^m$ fall into the mutually independent sets $x^1, \cdots, x^{m_1}; \cdots; x^{m_{p-1}+1}, \cdots, x^m$. Then the expected value of $(\lambda_I^{2/N})^h$, $h = \frac{1}{2}, 1, 1\frac{1}{2}, \cdots$, is greater under the null hypothesis H_I than under any alternative hypothesis in Ω.*

7. **The general regression problem.** Let the variates $x^1, \cdots, x^t$ be distributed according to the law

$$
(7.1) \qquad \frac{| B_{ij} |^{\frac{1}{2}}}{\pi^{\frac{1}{2}t}} e^{-B_{ij}(x^i - C^i_\mu x^\mu - C^i_\sigma x^\sigma)(x^j - C^j_\nu x^\nu - C^j_\tau x^\tau)}.
$$

Throughout this section, let the ranges of the indices be

$$
\begin{aligned}
i, j &= 1, \cdots, t & p, q &= t + 1, \cdots, m \\
r, s &= 1, \cdots, m & r', s' &= 1, \cdots, t + q \\
\mu, \nu &= t + 1, \cdots, t + q & \sigma, \tau &= t + q + 1, \cdots, m.
\end{aligned}
$$

In (7.1) we therefore have t random variates, and $m - t$ fixed variates. Consider the hypothesis H that the x^i are independent of the last set of x's, namely x^σ. We have

$$
\Omega\text{:} \quad \| B_{ij} \| \text{ positive definite, } -\infty < C^i_p < \infty,
$$

while for ω we impose the additional requirement

$$
C^i_\sigma = 0.
$$

Thus in general we have for the distribution of random samples 0_N, $N \geq m$,

$$
(7.2) \qquad \frac{| B_{ij} |^{\frac{1}{2}N}}{\pi^{\frac{1}{2}Nt}} e^{-\sum_{\alpha=1}^{N} B_{ij}(x^i_\alpha - C^i_p x^p_\alpha)(x^j_\alpha - C^j_q x^q_\alpha)}
$$

while when H is true, we have

$$P = \frac{|B_{ij}|^{\frac{1}{2}N}}{\pi^{\frac{1}{2}Nt}} e^{-\sum_{\alpha=1}^{N} B_{ij}(x_\alpha^i - C_\mu^i x_\alpha^\mu)(x_\alpha^j - C_\nu^j x_\alpha^\nu)} \tag{7.3}$$

Differentiating (7.2) with respect to the B's and C's and setting the derivatives equal to zero gives us the conditions

$$\sum_{\alpha=1}^{N} C_p^i x_\alpha^p x_\alpha^q = \sum_{\alpha=1}^{N} x_\alpha^i x_\alpha^q, \tag{7.4}$$

$$B^{ij} = \frac{2}{N} \sum_{\alpha=1}^{N} (x_\alpha^i - C_p^i x_\alpha^p)(x_\alpha^j - C_q^j x_\alpha^q). \tag{7.5}$$

As in section 2, we put

$$a^{rs} = \sum_{\alpha=1}^{N} x_\alpha^r x_\alpha^s.$$

and assume that the fixed values x_α^p have been so chosen that $\| a^{pq} \|$ is positive definite. Then (7.4) and (7.5) can be combined to give

$$B^{ij} = \frac{2}{N}(a^{ij} - a^{ip} a'_{pq} a^{qj}) = \frac{2}{N} \tilde{a}^{ij},$$

where $\| a'_{pq} \|^{-1} = \| a^{pq} \|$. It then follows that

$$P_\Omega = |\tilde{a}^{ij}|^{-\frac{1}{2}N} \left(\frac{N}{2\pi}\right)^{\frac{1}{2}Nt} e^{-\frac{1}{2}N}.$$

Similarly,

$$P_\omega = |\tilde{a}_0^{ij}|^{-\frac{1}{2}N} \left(\frac{N}{2\pi}\right)^{\frac{1}{2}Nt} e^{-\frac{1}{2}N},$$

where

$$\tilde{a}_0^{ij} = a^{ij} - a^{i\mu} a'_{\mu\nu} a^{\nu j}, \qquad \| a'_{\mu\nu} \|^{-1} = \| a^{\mu\nu} \|.$$

The matrix $\| a^{rs} \|$ will be positive definite except for a set of probability zero, so that we can consider $\| \tilde{a}^{ij} \|$ as the inverse of the matrix obtained by removing the last $m - t$ rows and columns of the inverse of $\| a^{rs} \|$, and $\| \tilde{a}_0^{ij} \|$ as the inverse of the matrix obtained by removing the last q rows and columns of $\| a^{r's'} \|^{-1}$. Then by the Jacobi theorem

$$|\tilde{a}^{ij}|^{-1} = \frac{|a^{pq}|}{|a^{rs}|}, \qquad |\tilde{a}_0^{ij}|^{-1} = \frac{|a^{\mu\nu}|}{|a^{r's'}|}$$

so that the appropriate likelihood ratio is given by

$$\lambda^{2/N} = \frac{|a^{rs}|}{|a^{r's'}|} \cdot \frac{|a^{\mu\nu}|}{|a^{pq}|}.$$

It will be advantageous to complete the matrix $\| B_{ij} \|$ in (7.1) by defining

$$
\begin{aligned}
B_{ip} &= -B_{ij}C^j_p, \\
B_{pq} &= C^i_p B_{ij} C^j_q.
\end{aligned} \tag{7.6}
$$

(Evidently $B_{ip} = 0$ for $i = 1, \cdots, t$ and fixed p, if and only if $C^j_p = 0$, $j = 1, \cdots, t$). We can now write (7.2) as

$$
P(x, B) = \frac{| B_{ij} |^{\frac{1}{2}N}}{\pi^{\frac{1}{2}Nt}} e^{-\sum_{\alpha=1}^{N} B_{ij}(x^i_\alpha + B^{ih}B_{hp}x^p_\alpha)(x^j_\alpha + B^{jk}B_{kq}x^q_\alpha)} \tag{7.7}
$$

We next notice that λ is invariant under the transformations

$$
x^i \to \alpha^i_j x^j, \qquad x^\sigma \to \beta^\sigma_\tau x^\tau,
$$

so that if we put

$$
I(B, \lambda_\epsilon) = \int_S P(x, B)\, dx^1_1 \cdots dx^t_N,
$$

where the integral is extended over the region

$$
S\colon \quad 0 \leq \lambda \leq \lambda_\epsilon,
$$

it turns out that

$$
I(B, \lambda_\epsilon) \equiv I(B^*, \lambda_\epsilon),
$$

provided

$$
B^*_{ij} = \alpha^k_i B_{kl} \alpha^l_j, \qquad B^*_{i\mu} = \alpha^k_i B_{k\mu}, \qquad B^*_{i\sigma} = \alpha^k_i B_{k\tau} \beta^\tau_\sigma.
$$

To prove the locally unbiased character of the test, we may therefore consider the derivatives

$$
\frac{\partial}{\partial B^*_{i\sigma}} I(B^*_0, \lambda_\epsilon), \qquad \frac{\partial^2}{\partial B^*_{i\sigma} \partial B^*_{j\tau}} I(B^*_0, \lambda_\epsilon),
$$

and assume that $\| B^*_{ij} \|$ and $\| a^{\mu\nu} \|$ are in diagonal form. We also observe that λ is unaltered by the transformation

$$
x^i \to x^i + B^{ik} B_{k\mu} x^\mu.
$$

We therefore have

$$
I(B^*, \lambda_\epsilon) = \frac{| B^*_{ij} |^{\frac{1}{2}N}}{\pi^{\frac{1}{2}Nt}} \int_S e^{-\sum_{\alpha=1}^{N} B^*_{ij}(x^i_\alpha + B^{*ik}B^*_{k\sigma}x^\sigma_\alpha)(x^j_\alpha + B^{*jl}B^*_{l\tau}x^\tau_\alpha)}\, dx.
$$

Thus,

$$
\frac{\partial}{\partial B^*_{k\sigma}} I(B^*_0, \lambda_\epsilon) = -2 \frac{| B^*_{ij0} |^{\frac{1}{2}N}}{\pi^{\frac{1}{2}Nt}} \int_S \sum_{\alpha=1}^{N} x^k_\alpha x^\sigma_\alpha e^{-\sum_{\alpha=1}^{N} B^*_{ij0} x^i_\alpha x^j_\alpha}\, dx,
$$

which is easily seen to be zero. Again, consider a non-repeated second partial derivative, say

$$\frac{\partial^2}{\partial B^*_{k\sigma}\partial B^*_{l\tau}} I(B^*_0, \lambda_\epsilon)$$

$$= -2\frac{|B^*_{ij}|^{\frac{1}{2}N}}{\pi^{\frac{1}{2}Nt}}\int_S \left(B^{*kl} a^{\sigma\tau} - 2\sum_{\alpha=1}^N x^\sigma_\alpha x^k_\alpha \cdot \sum_{\beta=1}^N x^\tau_\beta x^l_\beta\right) e^{-\sum_{\alpha=1}^N B^*_{ij} x^i_\alpha x^j_\alpha}\, dx.$$

This plainly vanishes if $k \neq l$; but it is by no means easy to see what happens when $k = l$, even when $\sigma \neq \tau$. Let us therefore study the distribution law of $\lambda^{2/N}$ for the case,

$$B_{i\sigma} = 0, \qquad i \neq 1.$$

(We shall not, however, assume that the transformation $B \to B^*$ has been made on the B's.)

Define

$$\tilde{B}_{pq} = B_{pq} - B_{pi}B^{ij}B_{jq},$$

$$\tilde{a}^{\sigma\tau} = a^{\sigma\tau} - a^{\sigma\mu}a_{\mu\nu}a^{\nu\tau},$$

where $\| a_{\mu\nu} \|$ now stands for the inverse of $\| a^{\mu\nu} \|$. These expressions will arise when we adapt Wilks' method of moment generating operators [13], based on the identity

$$\int e^{-B_{rs}a^{rs}} dx^1_1 \cdots dx^t_N = \pi^{\frac{1}{2}Nt} B^{-\frac{1}{2}N} \exp(-\tilde{B}_{pq}a^{pq}) \tag{7.8}$$

to the problem. We shall understand from now on that $B = |B_{ij}|$ and $\| B^{ij} \| = \| B_{ij} \|^{-1}$. Let us rearrange the form in the exponential on the right, thus:

$$\begin{aligned} \tilde{B}_{pq}a^{pq} &= (\tilde{B}_{\mu\nu}a^{\mu\nu} + 2B_{\mu\sigma}a^{\mu\sigma} + B_{\sigma\tau}a^{\sigma\tau} - 2B_{\mu i}B^{ij}B_{j\sigma}a^{\mu\sigma} \\ &\qquad - B_{\sigma i}B^{ij}B_{j\tau}a^{\sigma\mu}a_{\mu\nu}a^{\nu\tau}) - B_{\sigma i}B^{ij}B_{j\tau}\tilde{a}^{\sigma\tau} \\ &= Q - B_{\sigma i}B^{ij}B_{j\tau}\tilde{a}^{\sigma\tau} \\ &= Q - B^{ij}y_{ij}. \end{aligned}$$

A subscript β will denote the result of replacing $B_{r's'}$ by $B_{r's'} + \beta_{r's'}$, and a prime will indicate that each $\beta_{r's'}$ has been replaced by $\beta_{r's'} + \xi_{r'}\xi_{s'}$. Consider now the result of integrating the right hand side of (7.8) after these replacements have been made:

$$\begin{aligned} &\pi^{\frac{1}{2}Nt}\int B'^{-\frac{1}{2}N}_\beta \exp(-\tilde{B}'_{pq\beta}a^{pq})\, d\xi_1 \cdots d\xi_t\, d\xi_{t+1} \cdots d\xi_{t+q} \\ &\qquad = \pi^{\frac{1}{2}Nt}\int B'^{-\frac{1}{2}N}_\beta e^{B'^{ij}_\beta y_{ij}}\left(\int e^{-Q'_\beta}\, d\xi_\mu\right) d\xi_i, \end{aligned} \tag{7.9}$$

Let us integrate first with respect to the ξ_μ. Wilks has shown how to write Q'_β in the form

$$Q'_\beta = -Q'_{1\beta} + B_{pq\beta}a^{pq} + \frac{B_\beta}{B'_\beta}a^{\mu\nu}\xi_\mu\xi_\nu - 2B_{pi\beta}B'^{ij}_\beta a^{p\nu}\xi_j\xi_\nu,$$

where

$$Q'_{1\beta} = B_{\mu i\beta}B'^{ij}_\beta B_{j\nu\beta}a^{\mu\nu} + 2B_{\mu i\beta}B'^{ij}_\beta B_{j\sigma}a^{\mu\sigma} + B_{\sigma i}B'^{ij}_\beta B_{j\tau}a^{\sigma\mu}a_{\mu\nu}a^{\nu\tau}.$$

This latter expression is thus free of the ξ_μ. Consequently,

$$\int e^{-Q'_\beta}\,d\xi_\mu = \left(\frac{B'_\beta}{B_\beta}\right)^{\frac{1}{2}q} |\,a^{\mu\nu}\,|^{-\frac{1}{2}}\pi^{\frac{1}{2}q}e^{-B_{pq\beta}a^{pq}}e^{Q'_{1\beta}+Q'_{2\beta}},$$

where

$$Q'_{2\beta} = \frac{B'_\beta}{B_\beta}a_{\mu\nu}(a^{\mu p}B_{pi\beta}B'^{ij}_\beta\xi_j)(a^{\nu q}B_{qk\beta}B'^{kl}_\beta\xi_l),$$

which can be written

$$\frac{B'_\beta}{B_\beta}\{B_{\mu i\beta}B'^{ij}_\beta B_{\nu k\beta}B'^{kl}_\beta\xi_j\xi_l a^{\mu\nu} + 2B_{\mu i\beta}B'^{ij}_\beta B_{\sigma k}B'^{kl}_\beta\xi_j\xi_l a^{\mu\sigma}$$
$$+ B_{\sigma i}B'^{ij}_\beta B_{\tau k}B'^{kl}_\beta\xi_j\xi_l a^{\sigma\mu}a_{\mu\nu}a^{\nu\tau}\}.$$

The method of reduction used by Wilks can now be applied to $Q'_{1\beta}$ and $Q'_{2\beta}$, and gives

$$Q'_{1\beta} + Q'_{2\beta} = B_{\mu i\beta}B^{ij}_\beta B_{j\nu\beta}a^{\mu\nu} + 2B_{\mu i\beta}B^{ij}_\beta B_{j\sigma}a^{\mu\sigma} + B_{\sigma i}B^{ij}_\beta B_{j\tau}a^{\sigma\mu}a_{\mu\nu}a^{\nu\tau},$$

an expression which does not involve the ξ's. Thus

$$\int e^{-Q'_\beta}\,d\xi_\mu = \pi^{\frac{1}{2}q}|\,a^{\mu\nu}\,|^{-\frac{1}{2}}B_\beta^{-\frac{1}{2}q}e^{-Q_\beta}\cdot B'^{\frac{1}{2}q}_\beta. \tag{7.10}$$

Now the quantity

$$e^{B'^{ij}_\beta y_{ij}} = \sum_{\nu=0}^{\infty}\frac{(y_{ij}\bar{B}'^{ij}_\beta)^\nu}{\nu!}B'^{-\nu}_\beta,$$

where $\bar{B}^{ij}$ stands for the cofactor of B_{ij} in $||\,B_{ij}\,||$, can be expressed in terms of B'_β, provided we use our assumption that $B_{i\sigma} = 0$, $i \neq 1$, whereupon $y_{ij}B'^{ij}_\beta$ reduces to the single term yB'^{11}_β. In fact, we have

$$\begin{aligned} E[g_\beta|\,a^{rs}\,|^h] &= \bar{K}\prod_{i=1}^{t}\psi(N-m+t+1-i,2h)\,|\,a^{pq}\,|^h\pi^{\frac{1}{2}Nt}B_\beta^{-(\frac{1}{2}N+h)} \\ \times \exp(-\bar{B}_{pq\beta}a^{pq}) &= \bar{K}\prod_{i=1}^{j}\psi\cdot\pi^{\frac{1}{2}Nt}|\,a^{pq}\,|^hB_\beta^{-\frac{1}{2}q}e^{-Q_\beta}\sum_{\nu=0}^{\infty}\frac{(y\bar{B}^{11}_\beta)^\nu}{\nu!}B_\beta^{-[\frac{1}{2}(N-q)+h+\nu]}, \end{aligned} \tag{7.11}$$

where, following the notation used by Wilks [13],

$$g_\beta = e^{-\beta_{rs}a^{rs}}, \qquad \bar{K} = \pi^{-\frac{1}{2}Nt}B^{\frac{1}{2}N}\exp(-\bar{B}_{pq}a^{pq}),$$

$$\psi(a, b) = \frac{\Gamma[\frac{1}{2}(a+b)]}{\Gamma[\frac{1}{2}a]}.$$

And (7.11) can be written as

$$\begin{aligned} E[g_\beta \,|\, a^{rs} \,|^h] &= \bar{K}\pi^{\frac{1}{2}Nt}\prod_{i=1}^{t}\psi\cdot|\,a^{pq}\,|^h B_\beta^{-\frac{1}{2}q}e^{-Q_\beta} \\ &\quad \times \sum_{\nu=0}^{\infty}\frac{y^\nu}{\nu!}\frac{\Gamma[\frac{1}{2}(N-q)+h]}{\Gamma[\frac{1}{2}(N-q)+h+\nu]}\frac{\partial^\nu}{\partial u^\nu}(B_{\beta u}^{-[\frac{1}{2}(N-q)+h]})_{u=0}, \end{aligned} \tag{7.12}$$

where B_u stands for the result of replacing B_{11} by $B_{11} - u$. Changing $\beta_{r's'}$ into $\beta_{r's'} + \xi_{r'}\xi_{s'}$ and integrating, we then find that by virtue of (7.10)

$$\begin{aligned} E[g_\beta \,|\, a^{rs} \,|^h \,|\, a^{r's'} \,|^{-\frac{1}{2}}] &= \bar{K}\pi^{\frac{1}{2}Nt}\prod_{i=1}^{t}\psi\cdot|\,a^{pq}\,|^h\,|\,a^{\mu\nu}\,|^{-\frac{1}{2}}B_\beta^{-\frac{1}{2}q}e^{-Q_\beta} \\ &\quad \times \pi^{-\frac{1}{2}t}\prod_{\nu=0}^{\infty}\frac{y^\nu}{\nu!}\frac{\Gamma[\frac{1}{2}(N-q)+h]}{\Gamma[\frac{1}{2}(N-q)+h+\nu]}\frac{\partial^\nu}{\partial u^\nu}\int B_{\beta u}'^{-[\frac{1}{2}(N-q)+h]}\,d\xi_i\Big]_{u=0}. \end{aligned} \tag{7.13}$$

Now

$$\int B_{\beta u}'^{-[\frac{1}{2}(N-q)+h]}\,d\xi_i = B_{\beta u}^{-[\frac{1}{2}(N-1-q)+h]}\pi^{\frac{1}{2}t}\prod_{i=1}^{t}\psi(N-q+2h+1-i, -1),$$

so that (7.13) becomes

$$\begin{aligned} E[g_\beta \,|\, a^{rs} \,|^h \,|a^{r's'} \,|^{-\frac{1}{2}}] &= \bar{K}\pi^{\frac{1}{2}Nt}\prod_{i=1}^{t}\psi(N-m+t+1-i, 2h) \\ &\quad \times \prod_{i=1}^{t}\psi(N-q+2h+1-i, -1)\,|\,a^{pq}\,|^h\,|\,a^{\mu\nu}\,|^{-\frac{1}{2}} \\ &\quad \times \cdot B_\beta^{-\frac{1}{2}q}e^{-Q_\beta}\sum_{\nu=0}^{\infty}\frac{y^\nu}{\nu!}\frac{\Gamma[\frac{1}{2}(N-q)+h]}{\Gamma[\frac{1}{2}(N-q)+h+\nu]}\frac{\partial^\nu}{\partial u^\nu}(B_{\beta u}^{-[\frac{1}{2}(N-1-q)+h]})_{u=0}. \end{aligned} \tag{7.14}$$

Comparing (7.14) with (7.12), and making use of the fact that

$$\psi(a, -1)\psi(1-1, -1)\cdots\psi(a-2h+1, -1) = \psi(a, -2h),$$

we thus have

$$\begin{aligned} E[g_\beta \,|\, a^{rs} \,|^h \,|\, a^{r's'} \,|^{-h}] &= \bar{K}\pi^{\frac{1}{2}Nt}\prod_{i=1}^{t}\psi(N-m+t+1-i, 2h) \\ &\quad \times \prod_{i=1}^{t}\psi(N-q+2h+1-i, -2h)\,|\,a^{pq}\,|^h\,|\,a^{\mu\nu}\,|^{-h}B_\beta^{-\frac{1}{2}q}e^{-Q_\beta} \\ &\quad \times \sum_{\nu=0}^{\infty}\frac{y^\nu}{\nu!}\frac{\Gamma[\frac{1}{2}(N-q)+h]}{\Gamma[\frac{1}{2}(N-q)+h+\nu]}\frac{\partial^\nu}{\partial u^\nu}[B_{\beta u}^{-\frac{1}{2}(N-q)}]_{u=0}. \end{aligned}$$

Setting the β's equal to zero, performing the differentiation, and recalling the definitions of $\bar{K}$ and Q_β, we then find

$$
\begin{aligned}
E[(\lambda^{2/N})^h] &= \prod_{i=1}^{t} \psi(N - m + t + 1 - i, 2h) \prod_{i=1}^{t} \psi(N - q + 2h + 1 - i, -2h) \\
&\times e^{-yB^{11}} \sum_{\nu=0}^{\infty} \frac{(yB^{11})^\nu}{\nu!} \frac{\Gamma[\frac{1}{2}(N-q)+h]}{\Gamma[\frac{1}{2}(N-q)+h+\nu]} \frac{\Gamma[\frac{1}{2}(N-q)+\nu]}{\Gamma[\frac{1}{2}(N-q)]}.
\end{aligned}
\tag{7.15}
$$

Taking the first factor from each product, we can convert (7.15) into

$$
\begin{aligned}
&\prod_{i=2}^{t} \psi(N - m + t + 1 - i, 2h) \prod_{i=2}^{t} \psi(N - q + 2h + 1 - i, -2h) \\
&\qquad \times e^{-yB^{11}} \sum_{\nu=0}^{\infty} \frac{(yB^{11})^\nu}{\nu!} \frac{\Gamma[\frac{1}{2}(N-m+t)+h]}{\Gamma[\frac{1}{2}(N-m+t)]} \cdot \frac{\Gamma[\frac{1}{2}(N-q)+\nu]}{\Gamma[\frac{1}{2}(N-q)+h+\nu]}.
\end{aligned}
$$

This last product of ratios of Γ's is equivalent to

$$
\frac{\Gamma[\frac{1}{2}(N-q)+\nu]}{\Gamma[\frac{1}{2}(N-m+t)]\Gamma[\frac{1}{2}(m-t-q)+\nu]} \cdot \frac{\Gamma[\frac{1}{2}(m-t-q)+\nu]\Gamma[\frac{1}{2}(N-m+t)+h]}{\Gamma[\frac{1}{2}(N-q)+h+\nu]}.
$$

Thus the moments of $\lambda^{2/N}$ are connected with an integral equation of type B [12] and $\lambda^{2/N}$ is distributed like the product

$$
z \cdot \theta_2 \cdots \theta_t \qquad 0 \le z \le 1, 0 \le \theta_i \le 1,
$$

where the joint distribution of the θ's is

$$
\begin{aligned}
&f(\theta) \\
&= \prod_{i=2}^{t} \frac{\Gamma[\frac{1}{2}(N-q+1-i)]}{\Gamma[\frac{1}{2}(N-m+t+1-i)]\Gamma[\frac{1}{2}(m-t-q)]} \cdot \theta_i^{\frac{1}{2}(N-m+t+1-i)-1}(1-\theta_i)^{\frac{1}{2}(m-t-q)-1},
\end{aligned}
$$

and z is distributed independently of the θ's with the distribution

$$
F(z) = e^{-yB^{11}} \sum_{\nu=0}^{\infty} \frac{(yB^{11})^\nu}{\nu!} \frac{z^{\frac{1}{2}(N-m+t)-1}(1-z)^{\frac{1}{2}(m-t-q)+\nu-1}}{B[\frac{1}{2}(N-m+t), \frac{1}{2}(m-t-q)+\nu]}.
\tag{7.16}
$$

The probability that $0 \le \lambda \le \lambda_0$ is therefore

$$
I(y, \lambda_0) = \int_S f(\theta)F(z)\, dz\, d\theta_2 \cdots d\theta_t,
$$

where S is the region $0 \le \theta_2 \cdots \theta_t, z < \lambda_0^{2/N}$. Putting $\varphi(\theta)$ for the upper limit of z in S for fixed θ, and S_θ for the projection of S into the θ space, we then have

$$
I(y, \lambda_0) = \int_{S_\theta} f(\theta) \left\{ e^{-yB^{11}} \sum_{\nu=0}^{\infty} \frac{(yB^{11})^\nu}{\nu!} \int_0^\varphi \frac{z^{\frac{1}{2}(N-m+t)-1}(1-z)^{\frac{1}{2}(m-t-q)-1+\nu}}{B[\frac{1}{2}(N-m+t), \frac{1}{2}(m-t-q)+\nu]}\, dz \right\} d\theta.
$$

If we replace z by $(1 - z)$ we then find

$$
\begin{aligned}
I(y, \lambda_0) &= \int_{S_\theta} f(\theta) \\
&\times \left\{ e^{-yB^{11}} \sum_{\nu=0}^{\infty} \frac{(yB^{11})^\nu}{\nu!} \frac{B[\frac{1}{2}(m-t-q)+\nu, \frac{1}{2}(N-m+t); 1-\varphi]}{B[\frac{1}{2}(m-t-q)+\nu, \frac{1}{2}(N-m+t)]} \right\} d\theta.
\end{aligned}
\tag{7.17}
$$

As far as y is concerned, (7.17) is essentially the same as (2.8). The computation which was made there, together with the type of reasoning employed in the latter part of section 5 in connection with the independence test for several blocks, then shows that

$$\frac{\partial}{\partial y} I(y, \lambda_\epsilon) > 0 \qquad (0 < \epsilon < 1).$$

Remembering that

$$y = \tilde{a}^{\sigma\tau} B_{\sigma 1} B_{\tau 1},$$

we see that

$$\left(\frac{\partial y}{\partial B_{\sigma 1}}\right)_0 = 0, \qquad \frac{\partial^2 y}{\partial B_{\sigma 1}\, \partial B_{\tau 1}} = 2\tilde{a}^{\sigma\tau},$$

and we remark that the assumed positive definiteness of $\| a^{pq} \|$ implies that of $\| \tilde{a}^{\sigma\tau} \|$. Hence the relation

$$\left(\frac{\partial^2}{\partial B_{1\sigma}\, \partial B_{1\tau}} I(y, \lambda_\epsilon)\right)_0 = \left(\frac{\partial I}{\partial y}\right)_0 \tilde{a}^{\sigma\tau}$$

together with the fact that we could have obtained the analogue of (7.17) under the assumption

$$B_{i\sigma} = 0 \qquad i \neq i_0,$$

where i_0 is any fixed number in the set $1, \cdots, t$, shows that the matrix of second partial derivatives is positive definite when H is true.

Thus we have

THEOREM IV. *Let $x^1, \cdots, x^t$ be normally distributed about means which are linear functions of certain fixed variates $x^{t+1}, \cdots, x^m$. Then the likelihood ratio test for the hypothesis that the distribution of $x^1, \cdots, x^t$ depends only on a selected subset $x^{t+1}, \cdots, x^{t+q}$ of the fixed variates is locally unbiased.*

The result of this section has its most immediate application to those problems in the analysis of variance which require simultaneous consideration of several interrelated dependent variables $x^1, \cdots, x^t$ in conjunction with a given set of independent variables $x^{t+1}, \cdots, x^m$ [15]. For the usual hypothesis to be tested in this case is that $x^1, \cdots, x^t$ are jointly independent of, say, $x^{t+q+1}, \cdots, x^m$.

To return to the general case of (7.1), the method of this section can also be used to test the hypothesis that the regression coefficients referring to the x^σ have particular values, say

$$C^i_\sigma = C^i_{\sigma 0} \qquad i = 1, \cdots, t;\ \sigma = t + q + 1, \cdots, m,$$

the remaining C's and the B's being left unspecified. Since we have

$$x^i - C^i_\mu x^\mu - C^i_\sigma x^\sigma = x^i - C^i_\mu x^\mu - (C^i_\sigma - C^i_{\sigma 0}) x^\sigma - C^i_{\sigma 0} x^\sigma,$$

by the device of replacing x^i_α by $x^i_\alpha - C^i_{\sigma 0}x^\sigma_\alpha$, we can reduce this problem to that of testing the hypothesis that

$$C'^i_\sigma = C^i_\sigma - C^i_{\sigma 0} = 0.$$

Similarly, the problem of testing whether the linear functions $u^i_\sigma = \alpha^\tau_\sigma C^i_\tau$ have specified values $u^i_{\sigma 0}$ comes under the same heading [7].

A particularly interesting case of the general regression problem is that in which $m = t + q + 1$, so that the null hypothesis H states that the chance variables x^i are independent of the fixed variate x^m, though they may depend upon $x^{t+1}, \cdots, x^{m-1}$. In this case we are able to find the exact distribution law of $\lambda^{2/N}$ without assuming that any of the regression coefficients C^i are zero. For the quantity

$$\sum_{\nu=0}^{\infty} \frac{(y_{ij}\bar{B}^{ij})^\nu}{\nu!} B_\beta^{-[\frac{1}{2}(N-q)+h+\nu]}, \tag{7.18}$$

which would have occurred in (7.11) had it not been for the restriction $B_{i\sigma} = 0$ $(i \neq 1)$, can now be expressed in terms of B_β even without this restriction. By definition

$$y_{ij}\bar{B}^{ij} = \bar{a}^{mm}\bar{B}^{ij}B_{mi}B_{mj}$$

and the vanishing of the B_{mi} is equivalent to the vanishing of the regression coefficients C^i_m associated with x^m. And since

$$| B_{ij} - ua^{mm}B_{mi}B_{mj} | = B - ua^{mm}\bar{B}^{ij}B_{mi}B_{mj},$$

we can write (7.18) in the form

$$\sum_{\nu=0}^{\infty} \frac{1}{\nu!} \frac{\Gamma[\frac{1}{2}(N-q)+h]}{\Gamma[\frac{1}{2}(N-q)+h+\nu]} \cdot \frac{\partial^\nu}{\partial u^\nu} [B_{\beta u}^{-[\frac{1}{2}(N-q)+h]}]_{u=0},$$

where

$$|| B_{\beta u} || = || B_{ij\beta} - ua^{mm}B_{mi}B_{mj} ||$$

is positive definite provided u is sufficiently small. Thus the moments of $\lambda^{2/N}$ can be found from (7.15) if we put $a^{mm}B^{ij}B_{mi}B_{mj} = y_{ij}B^{ij}$ in place of yB^{11}. Moreover, it can be seen that when the value $m = t + q + 1$ is substituted into (7.15), that expression reduces to

$$E[(\lambda^{2/N})^h] = e^{-y_{ij}B^{ij}} \sum_{\nu=0}^{\infty} \frac{(y_{ij}B^{ij})^\nu}{\nu!} \frac{\mathrm{B}[\frac{1}{2}(N-m+1)+h, \frac{1}{2}(m-q-1)+\nu]}{\mathrm{B}[\frac{1}{2}(N-m+1), \frac{1}{2}(m-q-1)+\nu]}$$

so that $\lambda^{2/N}$ is distributed like w, where

$$f(w) = e^{-y_{ij}B^{ij}} \sum_{\nu=0}^{\infty} \frac{(y_{ij}B^{ij})^\nu}{\nu!} \frac{w^{\frac{1}{2}(N-m+1)-1}(1-w)^{\frac{1}{2}(m-q-1)-1+\nu}}{\mathrm{B}[\frac{1}{2}(N-m+1), \frac{1}{2}(m-q-1)+\nu]}. \tag{7.19}$$

The distribution law of $\lambda^{2/N}$ for this case is thus closely related to that obtained in the treatment of the regression problem with one dependent variate in section 2. Applying the argument used there, we can obtain:

THEOREM IVa. *The likelihood ratio test for the hypothesis that in a population of the type* (7.1) *the variates* x^i *are independent of* x^m—*the case* $m = t + q + 1$ *of Theorem IV—is completely unbiased.*

If we specialize the problem somewhat further, considering the case $q = 0$, $x_\alpha^m = 1$ (so that $m = t + 1$), we find that the likelihood ratio takes the form

$$\lambda^{2/N} = \frac{1}{1 + Nv_{ij}\bar{x}^i\bar{x}^j} = \frac{1}{1 + T},$$

where $v^{ij} = \sum_{\alpha=1}^{N} (x_\alpha^i - \bar{x}^i)(x_\alpha^j - \bar{x}^j)$, and T is Hotelling's generalization [5] of Student's ratio. In this case we are testing the hypothesis that the x^i are distributed with zero means. The exact distribution law of

$$T = \frac{1 - \lambda^{2/N}}{\lambda^{2/N}}$$

was recently published by P. L. Hsu [6], who obtained it in a very elegant fashion by means of the Laplace transform. He has also shown that the resulting test is *most powerful* in the sense that, of all critical regions S for which

$$P\{x \subset S\} = \epsilon + \tfrac{1}{2}\alpha B^{ij}b_ib_j + R(b)$$

(where ϵ and α are independent of the B^{ij} and of the means b_i, and R is an infinitesimal of at least the third order as all b_i tend to zero), the critical region defined by

$$S: \quad T \geq T_\epsilon$$

has the largest possible value of α. Tang's tables [11] make it evident that this largest possible value of α is actually positive and that the test is in fact unbiased for all values of the b's when $\epsilon = .05$ or $\epsilon = .01$. The results of this section may be used to show that this property extends to all probability levels other than $\epsilon = 0$ and $\epsilon = 1$.

The application of Hotelling's T is by no means confined to the above case. Other hypotheses which can be tested by means of this statistic are discussed by Hsu [6]. In addition it is now known that the Studentized D^2, devised by Mahalanobis for measuring the "distance" between two normal multivariate populations, is proportional to Hotelling's T. This fact is pointed out by R. C. Bose and N. Roy [1], who have obtained the exact distribution of D^2 for the case in which the two populations from which the samples are drawn are assumed to have the same matrix of variances and covariances, but are allowed to have different sets of means; their work, however, is quite independent of Hsu's. They also note that D^2 is proportional to the ratio which arises in Fisher's method of multiple measurements [4].

8. **Summary.** The method of likelihood-ratios is of practical as well as theoretical importance, because it provides a unified approach to the problem of testing statistical hypotheses. In this paper we have investigated many of the tests which this method yields when applied to hypotheses about sets of regression coefficients and covariances in normal populations. By studying the probability functions of the corresponding λ-criteria we are able to show that these tests are "good," in the sense that they are unbiased even for small samples.

Among the completely unbiased tests which can be based on the likelihood-ratio method, our discussion includes: the multiple correlation coefficient, with or without fixed variates [13]; Hotelling's generalized T test [6] and the statistically equivalent "Studentized D^2" [1]; the ordinary analysis of variance and covariance for orthogonal or non-orthogonal data [11, 16], as well as related tests of linear hypotheses in the case of one chance variable.

With respect to the analysis of variance for two or more variables [15] and certain other hypotheses regarding regression coefficients in multivariate populations, though there are indications that the tests are completely unbiased, we have succeeded in demonstrating this property only in the local sense.

Finally, the likelihood-ratio test for the hypothesis that the variates fall into certain specified mutually independent sets [14] is shown to be unbiased, at least locally, and has the additional property described in Theorem IIIa.

In conclusion, much more than a word of acknowledgment is due to Professor S. S. Wilks of Princeton University, to whom the writer is greatly indebted for advice and encouragement.

REFERENCES

[1] R. C. Bose and N. Roy, "The Distribution of the Studentized D²-statistic," *Sankhyā*, Vol. 4 (1938-39), pp. 19-38.

[2] G. W. Brown, "On the Power of the L_1 Test for Equality of Several Variances," *Annals of Math. Stat.*, Vol. 10 (1939), p. 124.

[3] R. A. Fisher, "The General Sampling Distribution of the Multiple Correlation Coefficient," *Proc. Roy. Soc. Lond.*, series A, Vol. 121 (1928), pp. 196-203.

[4] R. A. Fisher, "The Use of Multiple Measurements in Taxonomic Problems," *Annals of Eugenics*, Vol. 7 (1936), pp. 179-188.

[5] H. Hotelling, "The Generalization of Student's Ratio," *Annals of Math. Stat.*, Vol. 2 (1931), pp. 359-378.

[6] P. L. Hsu, "Notes on Hotelling's Generalized T," *Annals of Math. Stat.*, Vol. 9 (1938), pp. 231-243.

[7] S. Kolodziejczyk, "On an Important Class of Statistical Hypotheses," *Biometrika*, Vol. 27 (1935), pp. 161-190.

[8] J. Neyman and E. S. Pearson, "On the Use and Interpretation of Certain Test Criteria for Purposes of Statistical Inference," *Biometrika*, Vol. 20A (1928), pp. 175-240.

[9] J. Neyman and E. S. Pearson, "The Testing of Statistical Hypotheses in Relation to Probabilities a Priori," *Proc. Camb. Phil. Soc.*, Vol. 29 (1933), p. 493.

[10] J. Neyman and E. S. Pearson, "Contributions to the Theory of Testing Statistical Hypotheses," *Stat. Res. Mem.*, Vol. 1 (1936), pp. 1-37, Vol. 2 (1938), pp. 25-57.

[11] P. C. Tang, "The Power Function of the Analysis of Variance Tests with Tables and Illustrations of Their Use," *Stat. Res. Mem.*, Vol. 2 (1938), pp. 126-149 + viii.

[12] S. S. Wilks, "Certain Generalizations in the Analysis of Variance," *Biometrika*, Vol. 24 (1932), pp. 471–494.

[13] S. S. Wilks, "Moment-generating Operators for Determinants of Product Moments in Samples from a Norm: . System," *Ann. of Math.*, Vol. 35 (1934), pp. 312–340.

[14] S. S. Wilks, "On the Independence of k Sets of Normally Distributed Statistical Variables," *Econometrica*, Vol. 3 (1935), pp. 309–326.

[15] S. S. Wilks, "The Analysis of Variance for Two or More Variables," *Report of Third Annual Research Conference on Economics and Statistics*, (1937), pp. 82–85.

[16] S. S. Wilks, "The Analysis of Variance and Covariance in Non-Orthogonal Data," *Metron*, Vol. XIII (1938), pp. 141–154.

[17] J. Wishart, "The Generalized Product Moment Distribution in a Normal Distribution," *Biometrika*, Vol. 20^A (1928), pp. 32–52.

Princeton University,
Princeton, N. J.

STATISTICAL SEMINVARIANTS AND THEIR SETIMATES WITH PARTICULAR EMPHASIS ON THEIR RELATION TO ALGEBRAIC INVARIANTS

By Paul L. Dressel

TABLE OF CONTENTS

INTRODUCTION

An important portion of algebraic invariant theory has been that devoted to a certain class of invariants called seminvariants, semi-invariants, or more rarely, half-invariants. Of these terms, "seminvariant" seems to be the one now commonly accepted. The same three terms have been applied at various times and by various writers to a system of moment functions of importance in statistical theory. The statistician using these terms has frequently done so with an apology for appropriating a term of the algebraist. As a portion of this paper we shall show that the moment functions of this system are actually algebraic seminvariants, and that there are other systems of moment functions which are equally entitled to the name seminvariant.

The study of the statistical seminvariants of a population leads naturally to consideration of the problem of obtaining from a sample unbiased estimates of the value of these seminvariants. Estimates of this kind have been defined and computed by previous authors, but no simple method of obtaining the estimates has been given. In this paper a simple procedure for calculation is given and it is furthermore demonstrated that these estimates form an important phase of statistical seminvariant theory.

The system of notation used for moment functions is that of R. A. Fisher, although the actual letters used in representing particular moment functions are not altogether the same as those used by Fisher. In general, a moment function of the population has been indicated by a Greek letter, the corresponding sample moment function by the corresponding English letter and the estimate by the corresponding capital English letter.

A list of references appears at the end of the paper. Each reference has been assigned a number and this number placed in square brackets is used in the body of the paper to indicate the reference. Pages of the reference are indicated by additional numbers inserted in the parentheses and separated from the reference number by a semicolon.

I. THE RELATION OF THE ALGEBRAIC SEMINVARIANT THEORY TO THE MOMENT FUNCTIONS OF STATISTICS

The purposes of this chapter are: (1) to review briefly and give adequate references to certain important phases of algebraic seminvariant theory, (2) to apply this material to the moment functions of statistics.

1. **Definitions.** Any function of the coefficients of the binary form

$$f = \sum_{i=0}^{n} \binom{n}{i} a_i X^{n-i} Y^i, \qquad a_0 \neq 0, \tag{1}$$

which is invariant under the transformation

$$X = \gamma_1 \xi + \gamma_2 \eta, \qquad Y = \delta_1 \xi + \delta_2 \eta, \qquad \Delta = \begin{vmatrix} \gamma_1 & \gamma_2 \\ \delta_1 & \delta_2 \end{vmatrix} \neq 0, \tag{2}$$

is called an invariant of the form f. See Dickson [1; 31–36].

Any function of the coefficients of f which is invariant under the transformation

$$X = \xi + \gamma\eta, \qquad Y = \eta, \tag{3}$$

is called a seminvariant of f.

The two operators

$$\Omega = \sum_{i=1}^{n} i a_{i-1} \frac{\partial}{\partial a_i}, \qquad \bigcirc = \sum_{i=1}^{n} (n - i + 1) a_i \frac{\partial}{\partial a_{i-1}}, \tag{4}$$

are of fundamental importance in the theory of algebraic invariants and seminvariants and, indeed, invariants and seminvariants may be defined by means

of these operators. A necessary and sufficient condition that an homogeneous isobaric function of the coefficients of f be an invariant is that it be annihilated by both Ω and O. See Elliott [2; 113, 124]. The necessary and sufficient condition that an homogeneous isobaric function of the coefficients of f be a seminvariant is that it be annihilated by Ω. See Elliott [2; 127].

It should be noted that there is nothing in the definitions above which requires that invariants or seminvariants be integral, although usually only this type is discussed. In what follows we shall find it more profitable to discuss homogeneous isobaric fractional seminvariants, the fractional quality resulting from the appearance of a_0 in the denominator.

2. Complete Systems of Seminvariants. By direct application of the transformation (3) to f the system of seminvariants [1; 47]

$$A_r = \sum_{i=0}^{r} \binom{r}{i} \left(-\frac{a_1}{a_0}\right)^i \frac{a_{r-i}}{a_0}, \qquad r \leq n, \tag{5}$$

is obtained. This system is a complete system, [2; 44, 205, 206], in the sense that all other seminvariants fractional in a_0 and of degree 0 are expressible rationally and integrally in terms of this system.

Other such systems can be defined. The system of minimum degree seminvariants, the seminvariants of even weight being of degree 2 and those of odd weight being of degree 3, has played an important role in the algebraic seminvariant theory. Elliott [2; 207–209] discusses this system and gives the general formula for the even weight seminvariants of the system. So far as the present writer has been able to discover the general formula for the odd weight seminvariants has never been published, although Hammond [3] may have obtained it. After some lengthy but not difficult computation the result has been obtained, so that the last mentioned system of seminvariants is completely defined by

$$C_{2r} = \frac{1}{2} \sum_{i=0}^{2r} (-1)^i \binom{2r}{i} \frac{a_i a_{2r-i}}{a_0^2},$$

$$C_{2r+1} = \sum_{i=0}^{r} (-1)^{i+r} \binom{2r}{i+r} \frac{2i+1}{i+r+1} \frac{a_{r-i} a_{r+i+1}}{a_0^2} + \sum_{i=0}^{2r} (-1)^{i+1} \binom{2r}{i} \frac{a_1 a_i a_{2r-i}}{a_0^3}. \tag{6}$$

It is easily demonstrated that for each of the above seminvariants, and in fact for any seminvariant, the sum of the numerical coefficients is zero. Dickson [1; 55] gives a suggestion leading to a very simple proof.

3. The MacMahon Non-Unitary Symmetric Function Principle. Denoting the roots of $\sum_{i=0}^{n} \binom{n}{i} a_i X^{n-i} = 0$ by $\alpha_1, \alpha_2, \cdots \alpha_n$, the r-th power sum of these roots is defined by

$$s_r = \sum_{i=1}^{n} \alpha_i^r. \tag{7}$$

The form f may be written $\prod_{i=1}^{n} (X - \alpha_i Y)$.

By a result due to MacMahon [4; 131] the seminvariants of the form f are identical, except for numerical factors, with those symmetric functions of the roots of

$$g = \sum_{i=0}^{n} \frac{a_i}{i!} X^{n-i} = 0 \tag{8}$$

which when expressed in terms of sums of powers of these roots do not contain s_1. MacMahon called such symmetric functions "non-unitary."

As a result of this theorem, MacMahon was able to discuss the seminvariants of a binary form of infinite order by discussing the non-unitary symmetric functions of the roots of $\sum_{i=0}^{\infty} \frac{a_i}{i!} Y^i = 0$.

4. A Third Complete System of Seminvariants. By application of the result stated in the previous section, a third complete system of seminvariants can be immediately obtained. Obviously the power sums s_r, $r > 1$, are independent of s_1. By the Waring formula, Burnside and Panton [5; 91–92], if

$$\sum_{i=0}^{n} c_i Y^i = c_0 \prod_{i=0}^{n} (1 - \alpha_i Y)$$

then

$$s_r = \Sigma \frac{(-1)^\rho r(\rho - 1)!}{\pi_1! \pi_2! \cdots \pi_n!} \left(\frac{c_1}{c_0}\right)^{\pi_1} \left(\frac{c_2}{c_0}\right)^{\pi_2} \cdots \left(\frac{c_n}{c_0}\right)^{\pi_n}, \tag{9}$$

wherein

$$\rho = \sum_{i=1}^{n} \pi_i, \qquad r = \sum_{i=1}^{n} i\pi_i.$$

Then for

$$g = \sum_{i=0}^{n} \frac{a_i}{i!} X^{n-i},$$

$$-(r-1)! s_r = \Sigma \frac{(-1)^{\rho-1} r!(\rho - 1)! \left(\frac{a_1}{a_0}\right)^{\pi_1} \left(\frac{a_2}{a_0}\right)^{\pi_2} \cdots \left(\frac{a_n}{a_0}\right)^{\pi_n}}{\pi_1! \pi_2! \cdots \pi_n! (2!)^{\pi_2} \cdots (n!)^{\pi_n}}. \tag{10}$$

Placing $B_r = -(r - 1)! s_r$ the B's form a complete system of seminvariants. This result has some interesting statistical connections which will be mentioned later.

5. Linearly Independent Seminvariants. It follows from the MacMahon non-unitary symmetric function principle, or it can be proved easily in other ways, that the number of linearly independent seminvariants of a given weight r is

equal to the number of partitions of r which contain no unit part. Furthermore we have at our disposal a simple method for obtaining a set of linearly independent seminvariants of any given weight.

For many purposes the power product defined by Dwyer [6; 13] is more useful than the customary monomial symmetric function. The power product is defined by the right hand member and indicated by the left hand member of

$$(11) \qquad (q_1 \cdots q_r) = \sum_{i_1 \neq i_2 \neq \cdots \neq i_r} \alpha_{i_1}^{q_1} \alpha_{i_2}^{q_2} \cdots \alpha_{i_r}^{q_r},$$

where, for convenience, $q_1 \geq q_2 \geq \cdots \geq q_r$. The monomial symmetric function which will be denoted by $M(q_1 \cdots q_r)$ is related to the power product by the identity

$$(12) \qquad \pi_1! \cdots \pi_t! M(q_1^{\pi_1} q_2^{\pi_2} \cdots q_t^{\pi_t}) = (q_1^{\pi_1} q_2^{\pi_2} \cdots q_t^{\pi_t}),$$

so that a distinction occurs only when there are repeated exponents in the summation of (11).

If we desire a system of linearly independent seminvariants of weight 6, by the MacMahon principle we need only to compute the values of the power products (6), (42), (33), (222) in terms of the a's. In a somewhat different form these will be presented later.

6. **The Roberts Theorem.** Roberts, see [2; 231] and [5; 108], demonstrated the existence of a duality relationship between power sums, s's, and coefficients, a's such that corresponding to any seminvariant in terms of a's there exists a seminvariant in terms of s's obtained by replacing a_i by s_i. The proof consists of showing that the annihilator for seminvariants in terms of power sums is identical in form with Ω, a_i being replaced by s_i.

As a result of this duality, each of the systems of seminvariants which have been obtained yields, upon replacement of a_i by s_i, another system of seminvariants. In particular cases it may happen that the systems are identical when the identities connecting the a_i and s_i are taken into consideration.

We next wish to show that the systems of power sum seminvariants thus obtained either are identical with certain well known statistical moment functions or lead to new ones.

7. **Statistical Distributions Represented by Binary Forms.** The fact that statistical distributions may be represented by polynomials has long been recognized by statisticians, see Thiele [7; 24–26] and Bertilsen [8]. Indeed it was this fact which led Thiele to the definition of the seminvariants now called by his name. If we have given n observations $\alpha_1, \alpha_2, \cdots \alpha_n$, form the polynomial.

$$(13) \qquad F = \prod_{i=1}^{n} (X - \alpha_i) = \sum_{i=0}^{n} \binom{n}{i} \frac{a_i}{a_0} X^{n-i}.$$

F is not a binary form, but the seminvariant theory of binary forms is applicable since seminvariants are functions of the differences of the roots and are independent of the X and Y, which appear merely as convenient symbols to indicate the various terms of the algebraic form.

For distributions containing an infinite number of items the form F is of infinite order, but discussion of its seminvariants may be carried on by use of the MacMahon principle given in section 3.

8. **Three Systems of Statistical Seminvariants.** Before exhibiting some systems of statistical seminvariants it may be well to consider the meaning of "statistical seminvariant," for this phrase has been undefined. In fact the use of the phrase is merely a matter of convenience in that it emphasizes the fact that seminvariant moment functions have not previously been regarded as algebraic seminvariants. As used here a statistical seminvariant is an algebraic seminvariant which has some application in statistical theory.

The system of seminvariants (5) yields by application of the Roberts' Theorem the well known system of statistical seminvariants usually called central moments. If $\mu_r' = \frac{s_r}{n} = \frac{s_r}{s_0}$, the general formula may be written

$$\mu_r = \sum_{i=0}^{r} \binom{r}{i} \mu_{r-i}'(-\mu_1')^i. \tag{14}$$

The system of seminvariants (6) likewise leads to

$$\begin{aligned} \kappa_{2r} &= \frac{1}{2}\sum_{i=0}^{2r} (-1)^i \binom{2r}{i} \mu_i' \mu_{2r-i}', \\ \kappa_{2r+1} &= \sum_{i=0}^{r} (-1)^{i+r} \binom{2r}{i+r} \frac{2i+1}{i+r+1} \mu_{r-i}' \mu_{r+i+1}' \\ &\qquad + \sum_{i=0}^{2r} (-1)^{i+1} \binom{2r}{i} \mu_1' \mu_i' \mu_{2r-i}', \end{aligned} \tag{15}$$

a system which seems never to have been used by statisticians.

The system (10) leads to the well known Thiele seminvariants

$$\lambda_r = \Sigma \frac{(-1)^{\rho-1} r!(\rho - 1)!(\mu_1')^{\pi_1}(\mu_2')^{\pi_2} \cdots (\mu_r')^{\pi_r}}{\pi_1!\pi_2! \cdots \pi_r!(2!)^{\pi_2} \cdots (r!)^{\pi_r}}. \tag{16}$$

From sections 3 and 4 it is apparent that the general formula for the Thiele seminvariants is a special case of the Waring formula for power sums in terms of coefficients. It does not seem that this fact has been previously recognized. An equivalent way of stating this idea is to say that the Thiele seminvariant λ_r is, except for the factor $-(r-1)!$, the sum of the r-th powers of the roots of the equations obtained by setting the moment generating function,

$$M_x(Y) = \sum_{i=0}^{\infty} \mu_i' \frac{Y^i}{i!},$$

equal to zero.

It is of historical interest to note that MacMahon published his non-unitary function principle and the resulting set of seminvariants in 1884. Cayley [8] published an article in 1885 dealing with this same system. Roberts' Theorem having been known for some time (probably about 20 years), it seems probable that MacMahon and Cayley were aware of the Thiele seminvariants four to five years before Thiele's definition [9] by an entirely different method.

9. **Linearly Independent Statistical Seminvariants.** At the end of section 5 a method was indicated whereby a complete set of linearly independent seminvariants of a given weight r could be obtained. It has been noted previously that the one part symmetric function s_r or (r) leads to the Thiele seminvariant λ_r. As a further illustration consider the power product (22). From a table of symmetric functions we find that

$$(22) = \frac{2a_4}{4!a_0} - \frac{2a_3a_1}{3!a_0^2} + \frac{a_2^2}{2!2!a_0^2}$$
$$= \frac{2}{4!}\left(\frac{a_4}{a_0} - \frac{4a_3a_1}{a_0^2} + \frac{3a_2^2}{a_0^2}\right),$$

and by the Roberts' Theorem the statistical seminvariant

$$\frac{2}{4!}(\mu_4' - 4\mu_3'\mu_1' + 3\mu_2'^2)$$

is obtained. In similar fashion a system of linearly independent seminvariants of weight $\leqq 8$ have been computed and are given in Table I. For the sake of brevity they are expressed in terms of central moments. Hence the degree, by which is meant the maximum degree in the μ''s, is not apparent in the table. This definition of degree associates with the statistical seminvariant the degree (in the usual sense) of the corresponding homogeneous integral seminvariant.

10. **Statistical Invariants.** If the transformation

$$(17) \qquad x = \xi + mk\eta, \qquad y = m\eta$$

is applied to the binary form f and, if, in particular

$$k = -\frac{a_1}{a_0} \qquad \text{and} \qquad m = \left[\frac{a_2}{a_0} - \frac{a_1^2}{a_0^2}\right]^{-\frac{1}{2}}$$

one system of invariants of f under this transformation is found to be

$$(18) \qquad D_r = A_r/A_2^{\frac{1}{2}r}, \qquad r \leq n,$$

where A_r is defined in (5). By the Roberts Theorem we obtain the fact that the standard moment $\mu_r/\mu_2^{\frac{1}{2}r}$ is an invariant of f under this transformation. Thus the standard moments, or standard seminvariants in general, have also an algebraic connection. The effect of the transformation (17) on the roots of f is indicated by

$$x - \alpha_i y = \xi + mk\eta - m\alpha_i\eta = \xi - m(\alpha_i - k)\eta.$$

If m and k are defined as above, the result is the equivalent of measuring in standard units denoted by $\dfrac{\alpha_i - \mu_1'}{\sqrt{\mu_2}}$.

The system (18) is not a system of algebraic invariants, for algebraic invariants must be invariant under rotation, translation and change of scale, or stretching. The component parts of the above system are invariant only under the last two

TABLE I

Linearly Independent Seminvariants of Weight $\leqq 8$

Weight	Degree	Seminvariants
6	6	$\mu_6 - 15\mu_4\mu_2 - 10\mu_3^2 + 30\mu_2^3$
	4	$\mu_6 + 5\mu_4\mu_2 - 10\mu_3^2 - 30\mu_2^3$
	3	$\mu_6 - 15\mu_4\mu_2 + 20\mu_3^2 + 30\mu_2^3$
	2	$\mu_6 + 15\mu_4\mu_2 - 10\mu_3^2$
7	7	$\mu_7 - 21\mu_5\mu_2 - 35\mu_4\mu_3 + 56\mu_3\mu_2^2$
	5	$\mu_7 + 9\mu_5\mu_2 - 35\mu_4\mu_3 - 90\mu_3\mu_2^2$
	4	$\mu_7 - 21\mu_5\mu_2 + 25\mu_4\mu_3 + 30\mu_3\mu_2^2$
	3	$\mu_7 + 9\mu_5\mu_2 - 5\mu_4\mu_3$
8	8	$\mu_8 - 28\mu_6\mu_2 - 56\mu_5\mu_3 - 70\mu_4^2 + 210\mu_4\mu_2^2 + 280\mu_3^2\mu_2 - 105\mu_2^4$
	6	$\mu_8 + 14\mu_6\mu_2 - 56\mu_5\mu_3 - 35\mu_4^2 - 210\mu_4\mu_2^2 + 140\mu_3^2\mu_2 + 630\mu_2^4$
	5	$\mu_8 - 28\mu_6\mu_2 + 49\mu_5\mu_3 - 35\mu_4^2 + 420\mu_4\mu_2^2 - 490\mu_3^2\mu_2 - 630\mu_2^4$
	4	$\mu_8 - 28\mu_6\mu_2 - 56\mu_5\mu_3 + 105\mu_4^2 - 420\mu_4\mu_2^2 + 560\mu_3^2\mu_2 + 630\mu_2^4$
	4	$\mu_8 + 14\mu_6\mu_2 - 56\mu_5\mu_3 + 35\mu_4^2 - 210\mu_4\mu_2^2 + 140\mu_3^2\mu_2$
	3	$\mu_8 - 7\mu_6\mu_2 + 49\mu_5\mu_3 - 35\mu_4^2 + 105\mu_4\mu_2^2 - 70\mu_3^2\mu_2$
	2	$\mu_8 + 28\mu_6\mu_2 - 56\mu_5\mu_3 + 35\mu_4^2$

Weight	Degree	Seminvariants
0	1	μ_0
2	2	μ_2
3	3	μ_3
4	4	$\mu_4 - 3\mu_2^2$
	2	$\mu_4 + 3\mu_2^2$
5	5	$\mu_5 - 10\mu_3\mu_2$
	3	$\mu_5 + 2\mu_3\mu_2$

types of transformation. In statistics translation and change of scale ordinarily constitute the only desired transformations so that the standard seminvariants $\dfrac{\mu_r}{\mu_2^{\frac{1}{2}r}}, \dfrac{\lambda_r}{\lambda_2^{\frac{1}{2}r}}, \dfrac{\kappa_r}{\kappa_2^{\frac{1}{2}r}}, \cdots$ might well be called statistical invariants.

11. **Seminvariants and Invariants of Samples.** Consideration of the definition of seminvariants and invariants shows that:

1. A seminvariant is a seminvariant not because it is a function of deviations from the mean, but because it is a function of the differences of the observations;

2. An invariant is an invariant not because it is a seminvariant divided by the standard deviation raised to the proper power, but because it is a ratio of two seminvariants which are of the same order in powers of the observations.

These facts are important from the statistics viewpoint because they show that seminvariants and invariants of samples are also seminvariants and invariants of the population from which the samples are drawn.

II. ESTIMATES

1. **Power Product Seminvariants.** The Roberts Theorem set up a duality relationship between seminvariants expressed in terms of coefficients and seminvariants in terms of power sums. It can be shown that corresponding to each pair thus determined there exists a third seminvariant expressed in terms of power products. This leads to what may be called a triple system of seminvariants, the interrelationships being most apparent when all three seminvariants are expressed in terms of the notation defined by (11). The seminvariant $\frac{a_3}{a_0} - \frac{3a_2a_1}{a_0^2} + \frac{2a_1^3}{a_0^3}$ becomes in this notation

$$\frac{(111)}{n^{(3)}} - \frac{3(11)(1)}{n^{(2)}n} + \frac{2(1)^3}{n^3}.$$

The corresponding power sum seminvariant is

$$\frac{(3)}{n} - \frac{3(2)(1)}{n^2} + \frac{2(1)^3}{n^3},$$

while the power product seminvariant just mentioned is

$$\frac{(3)}{n} - \frac{3(21)}{n^{(2)}} + \frac{2(111)}{n^{(3)}}$$

The value of the power product notation lies in the fact that the numerical coefficients of the three seminvariants are then identical, while this is not the case when monomial and elementary symmetric functions are used.

Perhaps a few remarks are in order in regard to the proof of the relationship above expressed. The annihilator, corresponding to Ω, for seminvariants in terms of roots is, see [2; 230–31],

$$-D = \sum_{i=1}^{n} \frac{\partial}{\partial \alpha_i}.$$

It is easy to see that

$$-D\left[\frac{(p_1^{\pi_1}p_2^{\pi_2}\cdots p_s^{\pi_s})}{n^{(\rho)}}\right] = \frac{1}{n^{(\rho)}}\sum_{i=1}^{s}\pi_i p_i(p_1^{\pi_1}p_2^{\pi_2}\cdots p_i^{\pi_i-1}, p_i - 1, \cdot\cdot\, p_s^{\pi_s}),$$

and also that,

$$\frac{(p_1^{\pi_1}p_2^{\pi_2}\cdots p_{s-1}^{\pi_{s-1}}0)}{n^{(\rho)}} = \frac{(n-\rho+1)(p_1^{\pi_1}\cdots p_{s-1}^{\pi_{s-1}})}{n^{(\rho)}} = \frac{(p_1^{\pi_1}\cdots p_{s-1}^{\pi_{s-1}})}{n^{(\rho-1)}}.$$

Since

$$\Omega\left[\frac{(p_1)^{\pi_1}(p_2)^{\pi_2}\cdots(p_s)^{\pi_s}}{n^{\rho}}\right] = \frac{1}{n^{\rho}}\sum_{i=1}^{s}\pi_i p_i(p_1)^{\pi_1}(p_2)^{\pi_2}\cdots(p_i)^{\pi_i-1}(p_i-1)\cdots(p_s)^{\pi_s},$$

and

$$\frac{(p_1)^{\pi_1}(p_2)^{\pi_2}\cdots(p_{s-1})^{\pi_{s-1}}(0)}{n^{\rho}} = \frac{n(p_1)^{\pi_1}(p_2)^{\pi_2}\cdots(p_{s-1})^{\pi_{s-1}}}{n^{\rho}} = \frac{(p_1)^{\pi_1}\cdots(p_{s-1})^{\pi_{s-1}}}{n^{\rho-1}},$$

it becomes evident that corresponding to any power sum seminvariant there exists a power product seminvariant with the same numerical coefficients. The converse is also true.

2. **Unbiased Estimates of Rational Integral Moment Functions.** If τ represents a population parameter, and if t represents such a function of n observations that the expected value of t is equal to τ; then t is said to be an unbiased estimate of τ. See Tschuprow [11; 74–75], Bertilsen [8; 144], and Fisher [12].

Let $(p_1p_2\cdots p_s)$ denote a power product computed from a sample, the sample being from an infinite population. Then it is well known that

$$E\left[\frac{(p_1p_2\cdots p_s)}{n^{(s)}}\right] = \mu'_{p_1}\mu'_{p_2}\cdots\mu'_{p_s},$$

n being the number of items in the sample. If E^{-1} be interpreted as "unbiased estimate of," the above relation may also be written

$$E^{-1}[\mu'_{p_1}\mu'_{p_2}\cdots\mu'_{p_s}] = \frac{(p_1p_2\cdots p_s)}{n^{(s)}}, \tag{19}$$

and it is seen at once that the power product seminvariants defined in section 1, if computed from a sample of n observations, are the unbiased estimates of the corresponding power sum seminvariants of the infinite population from which the sample is drawn.

This provides an algebraic interpretation as well as a different approach to a topic which has already aroused considerable interest among statisticians. In 1927 Bertilsen [8; 144] gave the estimates of the first four Thiele seminvariants of the population in terms of Thiele seminvariants of the sample. In 1929 R. A. Fisher [12] also obtained these results and gave in addition the estimates of the fifth and sixth Thiele seminvariants. His results are in terms of sample moments. In 1937, P. S. Dwyer [13; 26] gave the estimates of the first five population central moments and indicated also means for obtaining the estimate of any rational integral isobaric moment function.

In the remainder of this chapter

(1) Dwyer's method will be extended and perhaps somewhat simplified,
(2) certain properties of this type of estimate will be pointed out,
(3) estimates of all seminvariants of weight ≤ 8 will be made available.

3. **Computation of Estimates.** From the relationship (19) it is possible to write down immediately in a simple, although not immediately useful, form the estimate of any rational integral moment function. Thus the fourth Thiele seminvariant λ_4 is given by

$$\lambda_4 = \mu'_4 - 4\mu'_3\mu'_1 - 3\mu'^2_2 + 12\mu'_2\mu'^2_1 - 6\mu'^4_1,$$

so that the estimate of λ_4 is

$$L_4 = \frac{(4)}{n} - \frac{4(31)}{n^{(2)}} - \frac{3(22)}{n^{(2)}} + \frac{12(211)}{n^{(3)}} - \frac{6(1111)}{n^{(4)}}.$$

Since power products are difficult to compute directly, it is necessary to express the estimates in terms of power sums. Dwyer [6; 30–33] gave a complete discussion of the problem of expanding power products in terms of power sums and also gave tables of power products in terms of power sums for weights $\leqq 6$. By use of (12) it is also possible to use tables giving monomial symmetric functions in terms of power sums. One table by J. R. Roe [14; plate 18] includes all cases of weight $\leqq 10$.

By use of such a table we find

$$\begin{aligned}
(31) &= -(4) + (3)(1), \\
(22) &= -(4) + (2)(2), \\
(211) &= 2(4) - 2(3)(1) - (2)(2) + (2)(1)^2, \\
(1111) &= -6(4) + 8(3)(1) + 3(2)(2) - 6(2)(1)^2 + (1)^4.
\end{aligned}$$

If these results are substituted in L_4 above and like terms are collected, it is found that

$$n^{(4)}L_4 = n^2(n+1)(4) - 4n(n+1)(3)(1) - 3n(n-1)(2)^2 + 12n(2)(1)^2 - 6(1)^4,$$

a result which agrees with that given by R. A. Fisher [12].

4. **The Dwyer Double Expansion Theorem.** The Dwyer double expansion theorem, [6; 34] and [11; 37–39], states that if any isobaric sum of power products of weight r indicated by

$$\text{(20)} \qquad \Sigma \frac{r!}{(q_1!)^{\pi_1} \cdots (q_t!)^{\pi_t} \pi_1! \cdots \pi_t!} b_{q_1^{\pi_1} \cdots q_t^{\pi_t}} (q_1^{\pi_1} \cdots q_t^{\pi_t})$$

be expanded in terms of power sums in a form indicated by

$$\text{(21)} \qquad \Sigma \frac{r!}{(p_1!)^{\pi_1} \cdots (p_s!)^{\pi_s} \pi_1! \cdots \pi_s!} a_{p_1^{\pi_1} \cdots p_s^{\pi_s}} (p_1)^{\pi_1} \cdots (p_s)^{\pi_s},$$

then the coefficient a_r of the power sum (r) is given by

$$\text{(22)} \qquad a_r = \Sigma(-1)^{\rho-1} \frac{(\rho-1)!\,r!}{(p_1!)^{\pi_1} \cdots (p_s!)^{\pi_s} \pi_1! \cdots \pi_s!} b_{p_1^{\pi_1}} \cdots b_{p_s^{\pi_s}},$$

and that the coefficient $a_{r_1 \cdots r_m}$ of $(r_1)(r_2) \cdots (r_m)$ is

$$\text{(23)} \qquad a_{r_1 \cdots r_m} = \overline{a_{r_1} a_{r_2} \cdots a_{r_m}}.$$

The barred product indicates a symbolic multiplication by suffixing of subscripts which is exemplified by

$$\overline{a_3 a_2} = \overline{(b_3 - 3b_{21} + 2b_{111})(b_2 - b_{11})} = b_{32} - b_{311} - 3b_{221} + 5b_{2111} - 2b_{11111} = a_{32}.$$

The application of this theorem to the present problem eliminates the use of tables and permits the independent computation of the coefficient of any particular products of power sums in the expansion in terms of power sums of any given estimate. The illustration given by Dwyer [13; 39, 40] exemplifies both of these points very well.

5. **Estimates of all Seminvariants of Weight** $\leqq 8$. If the estimates of any complete system of seminvariants and all products of these seminvariants up to and including weight r are known, then the estimates of all seminvariants of weight $\leqq r$ are obtainable as a linear combination of these known estimates. For example, suppose that we know the estimates of all Thiele seminvariants of weight $\leqq 5$ and wish to find the estimate of μ_5. Since $\mu_5 = \lambda_5 + 10\lambda_3\lambda_2$,

$$E^{-1}[\mu_5] = M_5 = E^{-1}[\lambda_5] + 10E^{-1}[\lambda_3\lambda_2] = L_5 + 10L_{32}.$$

In table II are given the estimates of all Thiele seminvariants and all products of Thiele seminvariants of weight $\leqq 8$. From this table the expressions for L_5 and L_{32} are obtained and, by taking the combination indicated above, it is seen that

$$\begin{aligned} n^{(5)}M_5 = (n^4 - 5n^3 + 10n^2)(5) - 5(n^3 - 5n^2 + 10n)(4)(1) \\ - 10(n^2 - n)(3)(2) + 10(n^2 - 4n + 8)(3)(1)^2 \\ + 30(n - 2)(2)^2(1) - 10n(2)(1)^3 + 4(1)^5, \end{aligned}$$

a result which checks with that given by Dwyer [13; 27]. In similar fashion the estimate of any other seminvariant of weight $\leqq 8$ can be obtained by use of table II.

6. **Computation Checks.** There are a number of checks which can be applied to the entries in table II. These may be of interest simply as properties of the estimates, and they may be of use in correcting errors which may possibly have crept into the tables.

When any power product of more than one part is expanded into power sums, the sum of the numerical coefficients of the expansion is zero. To prove this we need only to consider a set of observations of which one observation is unity and the rest are all zero. Then any power product of two or more parts is necessarily zero and all power sums are equal to unity. Hence the initial statement of the paragraph follows immediately.

From this fact it is apparent that the sum of the coefficients of L_r is $\frac{1}{n}$, and the sum of the coefficients of $L_{r_1r_2\cdots r_s}$ is zero. Thus for L_4 we have $\frac{n^3 + n^2 - 4(n^2 + n) - 3(n^2 - n) + 12n - 6}{n^{(4)}} = \frac{1}{n}$, and for L_{22} the sum of the coefficients is

$$\frac{1}{n^{(4)}}[-n^2 + n + 4n - 4 + n^2 - 3n + 3 - 2n + 1] = 0.$$

TABLE II

Estimates of All Thiele Seminvariants and Their Products of Weight ≤ 8

$w = 1$	nL_1
(1)	1

$w = 2$	$n^{(2)}L_2$
(2)	n
$(1)^2$	-1

$w = 3$	$n^{(3)}L_3$
(3)	n^2
(2)(1)	$-3n$
$(1)^3$	2

$w = 4$	$n^{(4)}L_4$	$n^{(4)}L_{22}$
(4)	$n^3 + n^2$	$-(n^2 - n)$
(3)(1)	$-4(n^2 + n)$	$4(n - 1)$
$(2)^2$	$-3(n^2 - n)$	$n^2 - 3n + 3$
$(2)(1)^2$	$12n$	$-2n$
$(1)^4$	-6	1

$w = 5$	$n^{(5)}L_5$	$n^{(5)}L_{32}$
(5)	$n^4 + 5n^3$	$-(n^3 - n^2)$
(4)(1)	$-5(n^3 + 5n^2)$	$5(n^2 - n)$
(3)(2)	$-10(n^3 - n^2)$	$n^3 - 2n^2 + 2n$
$(3)(1)^2$	$20(n^2 + 2n)$	$-(n^2 + 8n - 8)$
$(2)^2(1)$	$30(n^2 - n)$	$-3(n^2 - 2n + 2)$
$(2)(1)^3$	$-60n$	$+5n$
$(1)^5$	24	-2

$w = 6$	$n^{(6)}L_6$	$n^{(6)}L_{42}$	$n^{(6)}L_{33}$	$n^{(6)}L_{222}$
(6)	$n^5 + 16n^4 + 11n^3 - 4n^2$	$-(n^4 + 2n^3 - 7n^2 + 4n)$	$-(n^4 - 2n^3 + 5n^2 - 4n)$	$2(n^3 - 3n^2 + 2n)$
(5)(1)	$-6(n^4 + 16n^3 + 11n^2 - 4n)$	$6(n^3 + 2n^2 - 7n + 4)$	$6(n^3 - 2n^2 + 5n - 4)$	$-12(n^2 - 3n + 2)$
(4)(2)	$-15(n^4 - 4n^3 - n^2 + 4n)$	$n^4 - 10n^2 + 45n - 60$	$3(2n^3 - 5n^2 - 5n + 20)$	$-3(n^3 - 7n^2 + 20n - 20)$
$(3)^2$	$-10(n^4 - 2n^3 + 5n^2 - 4n)$	$2(2n^3 - 5n^2 - 5n + 20)$	$n^4 - 8n^3 + 25n^2 - 10n - 40$	$-2(3n^2 - 15n + 20)$
$(4)(1)^2$	$30(n^3 + 9n^2 + 2n)$	$-(n^3 + 15n^2 + 20n - 60)$	$-3(7n^2 - 15n + 20)$	$3(n^2 + 3n - 10)$
(3)(2)(1)	$120(n^3 - n)$	$-4(n^3 + 6n^2 - 25n + 30)$	$-6(n^3 - 4n^2 + 15n - 20)$	$12(n^2 - 4n + 5)$
$(2)^3$	$30(n^3 - 3n^2 + 2n)$	$-3(n^3 - 7n^2 + 20n - 20)$	$-3(3n^2 - 15n + 20)$	$n^3 - 9n^2 + 29n - 30$
$(3)(1)^3$	$-120(n^2 + 3n)$	$4(n^2 + 9n - 10)$	$4(n^2 + 3n)$	$-4(3n - 5)$
$(2)^2(1)^2$	$-270(n^2 - n)$	$3(5n^2 - 9n + 10)$	$9(n^2 - n)$	$-3(n^2 - 3n + 5)$
$(2)(1)^4$	$360n$	$-18n$	$-12n$	$3n$
$(1)^6$	-120	6	4	-1

TABLE II—*Continued*

$w = 7$	$n^{(7)}L_7$	$n^{(7)}L_{52}$	$n^{(7)}L_{43}$	$n^{(7)}L_{322}$
(7)	$n^6 + 42n^5 + 119n^4 - 42n^3$	$-(n^5 + 12n^4 - 31n^3 + 18n^2)$	$-(n^5 + 5n^3 - 6n^2)$	$2(n^4 - 3n^3 + 2n^2)$
(6)(1)	$-7(n^5 + 42n^4 + 119n^3 - 42n^2)$	$7(n^4 + 12n^3 - 31n^2 + 18n)$	$7(n^4 + 5n^2 - 6n)$	$-14(n^3 - 3n^2 + 2n)$
(5)(2)	$-21(n^5 + 12n^4 - 31n^3 + 18n^2)$	$n^5 + 6n^4 - 28n^3 + 99n^2 - 162n$	$3(3n^4 - 10n^3 + 5n^2 + 18n)$	$-2(n^4 - 6n^3 + 17n^2 - 18n)$
(4)(3)	$-35(n^5 + 5n^3 - 6n^2)$	$5(3n^4 - 10n^3 + 5n^2 + 18n)$	$n^5 - 6n^4 + 30n^3 - 35n^2 - 30n$	$-(n^4 - n^3 - 10n^2 + 20n)$
$(5)(1)^2$	$42(n^4 + 27n^3 + 44n^2 - 12n)$	$-(n^4 + 27n^3 + 224n^2 - 552n + 216)$	$-6(5n^3 - 5n^2 + 20n - 12)$	$2(n^3 + 15n^2 - 46n + 24)$
(4)(2)(1)	$210(n^4 + 6n^3 - 13n^2 + 6n)$	$-5(n^4 + 15n^3 - 58n^2 + 114n - 108)$	$-3(n^4 + 9n^3 - 20n^2 - 10n + 60)$	$13n^3 - 63n^2 + 140n - 120$
$(3)^2(1)$	$140(n^4 + 5n^2 - 6n)$	$-20(3n^3 - 10n^2 + 5n + 18)$	$-4(n^4 - 6n^3 + 30n^2 - 35n - 30)$	$4(n^3 - n^2 -$ [illegible] $+ 20)$
$(3)(2)^2$	$210(n^4 - 3n^3 + 2n^2)$	$-10(n^4 - 6n^3 + 17n^2 - 18n)$	$-3(n^4 - n^3 - 10n^2 + 20n)$	$n^4 - 7n^3 + 21n^2 - 20n$
$(4)(1)^3$	$-210(n^3 + 13n^2 + 6n)$	$5(n^3 + 19n^2 + 36n - 108)$	$2(n^3 + 34n^2 - 45n + 90)$	$-4(3n^2 + 2n - 15)$
$(3)(2)(1)^2$	$-1260(n^3 + n^2 - 2n)$	$30(n^3 + 7n^2 - 28n + 36)$	$12(2n^3 - 2n^2 + 15n - 30)$	$-2(n^3 + 12n^2 - 48n + 60)$
$(2)^3(1)$	$-630(n^3 - 3n^2 + 2n)$	$30(n^3 - 6n^2 + 17n - 18)$	$9(n^3 - n^2 - 10n + 20)$	$-3(n^3 - 7n^2 + 21n - 20)$
$(3)(1)^4$	$840(n^2 + 4n)$	$-20(n^2 + 10n - 12)$	$-14(n^2 + 4n)$	$n^2 + 24n - 40$
$(2)^2(1)^3$	$2520(n^2 - n)$	$-30(3n^2 - 5n + 6)$	$-42(n^2 - n)$	$2(4n^2 - 9n + 15)$
$(2)(1)^5$	$-2520n$	$84n$	$42n$	$-7n$
$(1)^7$	720	-24	-12	2

TABLE II—*Continued*

$w = 8$	$n^{(8)}L_8$	$n^{(8)}L_{62}$	$n^{(8)}L_{53}$
(8)	$n^7 + 99n^6 + 757n^5 + 141n^4 - 398n^3 + 120n^2$	$-(n^6 + 37n^5 - 39n^4 - 157n^3 + 278n^2 - 120n)$	$-(n^6 + 9n^5 - 23n^4 + 111n^3 - 218n^2 + 120n)$
(7)(1)	$-8(n^6 + 99n^5 + 757n^4 + 141n^3 - 398n^2 + 120n)$	$8(n^5 + 37n^4 - 39n^3 - 157n^2 + 278n - 120)$	$8(n^5 + 9n^4 - 23n^3 + 111n^2 - 218n + 120)$
(6)(2)	$-28(n^6 + 37n^5 - 39n^4 - 157n^3 + 278n^2 - 120n)$	$n^6 + 20n^5 + 3n^4 - 336n^3 + 1736n^2 - 4424n + 3360$	$13n^5 - 14n^4 - 57n^3 - 406n^2 + 2744n - 3360$
(5)(3)	$-56(n^6 + 9n^5 - 23n^4 + 111n^3 - 218n^2 + 120n)$	$2(13n^5 - 14n^4 - 57n^3 - 406n^2 + 2744n - 3360)$	$n^6 + n^5 - 51n^4 + 527n^3 - 1134n^2 - 2128n + 6720$
$(4)^2$	$-35(n^6 + n^5 + 33n^4 - 121n^3 + 206n^2 - 120n)$	$5(3n^5 - 11n^4 + 11n^3 + 119n^2 - 602n + 840)$	$5(n^5 + n^4 - 39n^3 + 119n^2 + 182n - 840)$
$(6)(1)^2$	$56(n^5 + 68n^4 + 359n^3 - 8n^2 - 60n)$	$-(n^5 + 48n^4 + 1039n^3 - 1428n^2 - 2660n + 3360)$	$-(41n^4 + 238n^3 - 701n^2 + 2702n - 3360)$
(5)(2)(1)	$336(n^5 + 23n^4 - 31n^3 - 23n^2 + 30n)$	$-6(n^5 + 33n^4 - 11n^3 - 393n^2 + 1330n - 1680)$	$-3(n^5 + 27n^4 - 79n^3 + 413n^2 - 1946n + 3360)$
(4)(3)(1)	$560(n^5 + 5n^4 + 5n^3 - 5n^2 - 6n)$	$-10(25n^4 - 58n^3 - 13n^2 + 70n + 336)$	$-5(n^5 + 9n^4 - 43n^3 + 215n^2 - 182n - 672)$
$(4)(2)^2$	$420(n^5 + 2n^4 - 25n^3 + 46n^2 - 24n)$	$-15(n^5 - 2n^4 - 27n^3 + 236n^2 - 700n + 672)$	$-15(5n^4 - 26n^3 - 25n^2 + 406n - 672)$
$(3)^2(2)$	$560(n^5 - 4n^4 + 11n^3 - 20n^2 + 12n)$	$-10(n^5 - 5n^3 - 96n^2 + 532n - 672)$	$-10(n^5 - 11n^4 + 61n^3 - 101n^2 - 238n + 672)$
$(5)(1)^3$	$-336(n^4 + 38n^3 + 99n^2 - 18n)$	$6(n^4 + 38n^3 + 339n^2 - 738n)$	$2(n^4 + 68n^3 + 159n^2 - 288n + 756)$
$(4)(2)(1)^2$	$-2520(n^4 + 10n^3 - 17n^2 + 6n)$	$45(n^4 + 18n^3 - 41n^2 + 22n)$	$15(n^4 + 28n^3 - 113n^2 + 264n - 252)$
$(3)^2(1)^2$	$-1680(n^4 + 2n^3 + 7n^2 - 10n)$	$10(n^4 + 50n^3 - 121n^2 - 122n + 672)$	$20(n^4 - n^3 + 9n^2 + 57n - 210)$
$(3)(2)^2(1)$	$-5040(n^4 - 2n^3 - n^2 + 2n)$	$120(n^4 - n^3 - 16n^2 + 70n - 84)$	$60(n^4 - 6n^3 + 35n^2 - 126n + 168)$
$(2)^4$	$-630(n^4 - 6n^3 + 11n^2 - 6n)$	$30(n^4 - 12n^3 + 62n^2 - 147n + 126)$	$90(n^3 - 10n^2 + 35n - 42)$
$(4)(1)^4$	$1680(n^3 + 17n^2 + 12n)$	$-30(n^3 + 23n^2 + 54n - 168)$	$-10(n^3 + 38n^2 - 45n + 126)$
$(3)(2)(1)^3$	$13440(n^3 + 2n^2 - 3n)$	$-240(n^3 + 8n^2 - 31n + 42)$	$-20(7n^3 + 2n^2 + 45n - 126)$
$(2)^3(1)^2$	$10080(n^3 - 3n^2 + 2n)$	$-60(5n^3 - 27n^2 + 76n - 84)$	$-90(n^3 - 2n^2 - 5n + 14)$
$(3)(1)^5$	$-6720(n^2 + 5n)$	$120(n^2 + 11n - 14)$	$64(n^2 + 5n)$
$(2)^2(1)^4$	$-25200(n^2 - n)$	$90(7n^2 - 11n + 14)$	$240(n^2 - n)$
$(2)(1)^6$	$20160n$	$-480n$	$-192n$
$(1)^8$	-5040	120	48

TABLE II—*Continued*

$w = 8$	$n^{(8)}L_{44}$	$n^{(8)}L_{422}$	$n^{(8)}L_{332}$	$n^{(8)}L_{2222}$
(8)	$-(n^6 + n^5 + 33n^4 - 121n^3 + 206n^2 - 120n)$	$2(n^5 + 2n^4 - 25n^3 + 46n^2 - 24n)$	$2(n^5 - 4n^4 + 11n^3 - 20n^2 + 12n)$	$-6(n^4 - 6n^3 + 11n^2 - 6n)$
(7)(1)	$8(n^5 + n^4 + 33n^3 - 121n^2 + 206n - 120)$	$-16(n^4 + 2n^3 - 25n^2 + 46n - 24)$	$-16(n^4 - 4n^3 + 11n^2 - 20n + 12)$	$48(n^3 - 6n^2 + 11n - 6)$
(6)(2)	$4(3n^5 - 11n^4 + 11n^3 + 119n^2 - 602n + 840)$	$-2(n^5 - 2n^4 - 27n^3 + 236n^2 - 700n + 672)$	$-(n^5 - 5n^3 - 96n^2 + 532n - 672)$	$8(n^4 - 12n^3 + 62n^2 - 147n + 126)$
(5)(3)	$8(n^5 + n^4 - 39n^3 + 119n^2 + 182n - 840)$	$-4(5n^4 - 26n^3 - 25n^2 + 406n - 672)$	$-2(n^5 - 11n^4 + 61n^3 - 101n^2 - 238n + 672)$	$48(n^3 - 10n^2 + 35n - 42)$
$(4)^2$	$n^6 - 13n^5 + 76n^4 - 37n^3 - 497n^2 - 490n + 4200$	$(-n^5 - 8n^4 + 33n^3 + 62n^2 - 868n + 1680)$	$-2(4n^4 - 31n^3 + 65n^2 + 112n - 420)$	$3(n^4 - 14n^3 +$ [illegible]$5n^2 - 322n + 420)$
$(6)(1)^2$	$-8(5n^4 - 2n^3 + 121n^2 - 364n + 420)$	$2(n^4 + 26n^3 + 29n^2 - 464n + 588)$	$n^4 + 56n^3 - 229n^2 + 520n - 588$	$-8(n^3 + 9n^2 - 64n + 84)$
(5)(2)(1)	$-48(2n^4 - 5n^3 - 14n^2 + 119n - 210)$	$12(n^4 + 3n^3 - 53n^2 + 211n - 294)$	$6(2n^4 - 11n^3 + 56n^2 - 197n + 294)$	$-48(n^3 - 9n^2 + 32n - 42)$
(4)(3)(1)	$-8(n^5 - 8n^4 + 81n^3 - 232n^2 + 98n + 420)$	$4(2n^4 + 9n^3 - 64n^2 - n + 294)$	$2(5n^4 - 23n^3 + 57n^2 + 15n - 294)$	$-24(n^3 - 4n^2 - 5n + 28)$
$(4)(2)^2$	$-6(n^5 - 8n^4 + 33n^3 + 62n^2 - 868n + 1680)$	$n^5 - n^4 - 94n^3 + 868n^2 - 2952n + 3528$	$3(2n^4 - 7n^3 - 57n^2 + 380n - 588)$	$-6(n^4 - 16n^3 + 104n^2 - 305n + 336)$
$(3)^2(2)$	$-16(4n^4 - 31n^3 + 65n^2 + 112n + 420)$	$4(2n^4 - 7n^3 - 57n^2 + 380n - 588)$	$n^5 - 12n^4 + 63n^3 - 52n^2 - 536n + 1176$	$-8(3n^3 - 33n^2 + 128n - 168)$
$(5)(1)^3$	$48(3n^3 - 4n^2 + 31n - 42)$	$-4(3n^3 + 32n^2 - 77n - 42)$	$-2(5n^3 + 34n^2 - 117n + 126)$	$48(n^2 - 3n)$
$(4)(2)(1)^2$	$24(n^4 + 2n^3 + 32n^2 - 155n + 210)$	$-2(n^4 + 20n^3 - 22n^2 - 119n + 210)$	$-3(19n^3 - 92n^2 + 223n - 210)$	$12(n^3 - 3n^2 + 2n)$
$(3)^2(1)^2$	$16(n^4 - 4n^3 + 50n^2 - 167n + 210)$	$-4(6n^3 + 11n^2 - 185n + 378)$	$-(n^4 + 8n^3 - 29n^2 + 176n - 476)$	$8(9n^2 - 57n + 98)$
$(3)(2)^2(1)$	$24(n^4 + 8n^3 - 91n^2 + 322n - 420)$	$-4(n^4 + 11n^3 - 136n^2 + 526n - 672)$	$-6(n^4 - 8n^3 + 49n^2 - 166n + 224)$	$24(n^3 - 10n^2 + 38n - 49)$

$(2)^4$	$9(n^4 - 14n^3 + 95n^2 - 322n + 420)$	$-3(n^4 - 16n^3 + 104n^2 - 305n + 336)$	$-3(3n^3 - 33n^2 + 128n - 168)$	$n^4 - 18n^3 + 125n^2 - 384n + 441$
$(4)(1)^4$	$-12(n^3 + 17n^2 + 12n)$	$n^3 + 35n^2 + 138n - 504$	$41n^2 - 53n + 42$	$-6(n^2 + 7n - 28)$
$(3)(2)(1)^3$	$-96(n^3 + 2n^2 - 3n)$	$8(n^3 + 20n^2 - 87n + 126)$	$2(5n^3 + 14n^2 - 17n - 42)$	$-16(3n^2 - 14n + 21)$
$(2)^3(1)^2$	$-72(n^3 - 3n^2 + 2n)$	$18(n^3 - 7n^2 + 24n - 28)$	$3(3n^3 - 12n^2 + 15n + 14)$	$-4(n^3 - 9n^2 + 35n - 42)$
$(3)(1)^5$	$48(n^2 + 5n)$	$-4(n^2 + 23n - 42)$	$-4(n^2 + 11n - 14)$	$8(3n - 7)$
$(2)^2(1)^4$	$180(n^2 - n)$	$-3(11n^2 - 23n + 42)$	$-3(7n^2 - 11n + 14)$	$6(n^2 - 3n + 7)$
$(2)(1)^6$	$-144n$	$24n$	$16n$	$-4n$
$(1)^8$	36	-6	-4	1

A condition satisfied by the coefficients of any seminvariant is that their sum is equal to zero (See section 2). This provides another check on the entries of table II, although the seminvariant must be written in homogeneous form before the check is applied. Thus we may write

$$L_4 = \frac{n^3}{n^{(4)}}\left[(n+1)\frac{(4)}{n} - 4(n+1)\frac{(3)(1)}{n^2} - 3(n-1)\frac{(2)^2}{n^2} + 12n\frac{(2)(1)^2}{n^3} - 6n\frac{(1)^4}{n^4}\right],$$

and the sum of coefficients is

$$(n+1) - 4(n+1) - 3(n-1) + 12n - 6n = 0.$$

Several checks arise from the fact (see section 6) that every seminvariant must be annihilated by the operator

$$\Omega' = \sum_{i=1}^{n} i s_{i-1} \frac{\partial}{\partial s_i}. \tag{24}$$

Another check results from the discussion of the next section and is so apparent as to need no comment.

All the checks mentioned in this section are applicable to the estimate of any seminvariant.

7. **Estimates as Sums of Simple Seminvariants.** A seminvariant such as L_4 in which the coefficients of the m''s are functions of n will be called a composite seminvariant, while a seminvariant in which the coefficients of the m''s are purely numerical will be called simple. The fact that is to be established in this section is that every composite seminvariant is the sum of simple seminvariants. As an illustration consider L_4. It is apparent that

$$L_4 = \frac{n^4}{n^{(4)}} l_4 + \frac{n^3}{n^{(4)}} k_4,$$

where l_4 and k_4 are seminvariants of the sample corresponding to λ_4 and κ_4. Both l_4 and k_4 are simple seminvariants.

That a composite seminvariant may always be expressed as a sum of simple seminvariants can be demonstrated by considering the effect of Ω', (24), on a composite seminvariant. The coefficients are polynomials in n and are unaffected by the operator. The expression resulting from application of the operator can vanish only if the coefficient of n^r vanishes for every r. Thus a composite seminvariant which has r different powers of n appearing in its coefficients is expressible as the sum of r simple seminvariants, which are not necessarily distinct. Table III exhibits the estimates of Thiele seminvariants of weight $\leqq 6$ as sums of simple seminvariants.

Since the factors, appearing in front of each of the simple seminvariants in the expression resulting from breaking down a composite seminvariant, are of

TABLE III

Estimates as Sums of Seminvariants

L_2	$\frac{n^2}{n^{(2)}}$
$\frac{(2)}{n}$	1
$\frac{(1)^2}{n^2}$	−1

L_3	$\frac{n^3}{n^{(3)}}$
$\frac{(3)}{n}$	1
$\frac{(2)(1)}{n^2}$	−3
$\frac{(1)^3}{n^3}$	2

L^4	$\frac{n^4}{n^{(4)}}$	$\frac{n^3}{n^{(4)}}$
$\frac{(4)}{n}$	1	1
$\frac{(3)(1)}{n^2}$	−4	−4
$\frac{(2)^2}{n^2}$	−3	3
$\frac{(2)(1)^2}{n^3}$	12	
$\frac{(1)^4}{n^4}$	−6	

L_{22}	$\frac{n^4}{n^{(4)}}$	$-\frac{n^3}{n^{(4)}}$	$\frac{n^2}{n^{(4)}}$
$\frac{(4)}{n}$		1	1
$\frac{(3)(1)}{n^2}$		−4	−4
$\frac{(2)^2}{n^2}$	1	3	3
$\frac{(2)(1)^2}{n^3}$	−2		
$\frac{(1)^4}{n^4}$	1		

L_{32}	$\frac{n^5}{n^{(5)}}$	$-\frac{n^4}{n^{(5)}}$	$\frac{n^3}{n^{(5)}}$
$\frac{(5)}{n}$		1	1
$\frac{(4)(1)}{n^2}$		−5	−5
$\frac{(3)(2)}{n^2}$	1	2	2
$\frac{(3)(1)^2}{n^3}$	−1	8	8
$\frac{(2)^2(1)}{n^3}$	−3	−6	−6
$\frac{(2)(1)^3}{n^4}$	5		
$\frac{(1)^5}{n^5}$	−2		

L_6	$\frac{n^6}{n^{(6)}}$	$\frac{2n^5}{n^{(6)}}$	$\frac{n^4}{n^{(6)}}$	$-\frac{4n^3}{n^{(6)}}$
$\frac{(6)}{n}$	1	8	11	1
$\frac{(5)(1)}{n^2}$	−6	−48	−66	−6
$\frac{(4)(2)}{n^2}$	−15	−15	105	15
$\frac{(3)^2}{n^2}$	−10	10	−50	−10
$\frac{(4)(1)^2}{n^3}$	30	135	60	
$\frac{(3)(2)(1)}{n^3}$	120		−120	
$\frac{(2)^3}{n^3}$	30	−45	60	
$\frac{(3)(1)^3}{n^4}$	−120	−180		
$\frac{(2)^2(1)^2}{n^4}$	−270	135		
$\frac{(2)(1)^4}{n^5}$	360			
$\frac{(1)^6}{n^6}$	−120			

L_5	$\frac{n^5}{n^{(5)}}$	$\frac{5n^4}{n^{(5)}}$
$\frac{(5)}{n}$	1	1
$\frac{(4)(1)}{n^2}$	−5	−5
$\frac{(3)(2)}{n^2}$	−10	2
$\frac{(3)(1)^2}{n^3}$	20	8
$\frac{(2)^2(1)}{n^3}$	30	−6
$\frac{(2)(1)^3}{n^4}$	−60	
$\frac{(1)^5}{n^5}$	24	

TABLE III—*Continued*

L_{42}	$\frac{n^6}{n^{(6)}}$	$-\frac{n^5}{n^{(6)}}$	$-\frac{2n^4}{n^{(6)}}$	$\frac{n^3}{n^{(6)}}$	$-\frac{4n^2}{n^{(6)}}$
$\frac{(6)}{n}$		1	1	7	1
$\frac{(5)(1)}{n^2}$		−6	−6	−42	−6
$\frac{(4)(2)}{n^2}$	1	15	5	45	15
$\frac{(3)^2}{n^2}$		−4	5	−10	−10
$\frac{(4)(1)^2}{n^3}$	−1		10	60	
$\frac{(3)(2)(1)}{n^3}$	−4	24	−50	−120	
$\frac{(2)^3}{n^3}$	−3	−21	30	60	
$\frac{(3)(1)^3}{n^4}$	4	−36	20		
$\frac{(2)^2(1)^2}{n^4}$	15	27	−15		
$\frac{(2)(1)^4}{n^5}$	−18				
$\frac{(1)^6}{n^6}$	6				

L_{33}	$\frac{n^6}{n^{(6)}}$	$-\frac{n^5}{n^{(6)}}$	$\frac{n^4}{n^{(6)}}$	$-\frac{5n^3}{n^{(6)}}$	$\frac{4n^2}{n^{(6)}}$
$\frac{(6)}{n}$		1	2	1	1
$\frac{(5)(1)}{n^2}$		−6	−12	−6	−6
$\frac{(4)(2)}{n^2}$		−6	−15	3	15
$\frac{(3)^2}{n^2}$	1	8	25	2	−10
$\frac{(4)(1)^2}{n^3}$		21	45	12	
$\frac{(3)(2)(1)}{n^3}$	−6	−24	−90	−24	
$\frac{(2)^3}{n^3}$		9	45	12	
$\frac{(3)(1)^3}{n^4}$	4	−12			
$\frac{(2)^2(1)^2}{n^4}$	9	9			
$\frac{(2)(1)^4}{n^5}$	−12				
$\frac{(1)^6}{n^6}$	4				

L_{222}	$\frac{n^6}{n^{(6)}}$	$-\frac{3n^5}{n^{(6)}}$	$\frac{n^4}{n^{(6)}}$	$-\frac{6n^3}{n^{(6)}}$	$\frac{4n^2}{n^{(6)}}$
$\frac{(6)}{n}$			2	1	1
$\frac{(5)(1)}{n^2}$			−12	−6	−6
$\frac{(4)(2)}{n^2}$		1	21	10	15
$\frac{(3)^2}{n^2}$			−6	−5	−10
$\frac{(4)(1)^2}{n^3}$		−1	9	5	
$\frac{(3)(2)(1)}{n^3}$		−4	−48	−10	
$\frac{(2)^3}{n^3}$	1	3	29	5	
$\frac{(3)(1)^3}{n^4}$		4	20		
$\frac{(2)^2(1)^2}{n^4}$	−3	−3	−15		
$\frac{(2)(1)^4}{n^5}$	3				
$\frac{(1)^6}{n^6}$	−1				

successively lower order with respect to n; it is possible to obtain approximations of various orders to the value of an estimate by using the appropriate portion of the expression given in the table.

8. **The Estimates of the κ's.** The seminvariant κ_r possesses an interesting property which will be called invariance under estimate. By this is meant that the estimate of κ_r is k_r multiplied by a suitable factor. In particular, $\kappa_2 = \mu_2$ and $\kappa_3 = \mu_3$ and it is well known that

$$E^{-1}[\mu_2] = \frac{n^2}{n^{(2)}} m_2, \qquad E^{-1}[\mu_3] = \frac{n^3}{n^{(3)}} m_3$$

so that the κ_r certainly possesses the property for $r = 2$ and 3. It can be shown, however, that

$$K_{2r} = \frac{n^2}{n^{(2)}} K_{2r}, \qquad K_{2r+1} = \frac{n^3}{n^{(3)}} K_{2r+1}. \tag{25}$$

From (15)

$$\kappa_{2r} = \frac{1}{2} \sum_{i=0}^{2r} \binom{2r}{i} \mu_i' \mu_{2r-i}'$$

so that

$$K_{2r} = \frac{1}{2} \sum_{i=1}^{2r-1} (-1)^i \frac{(i, 2r - i)}{n^{(2)}} + \frac{(2r)}{n}.$$

By the Binet-Waring identities [15; 6–7]

$$(a \cdot b) = (a)(b) - (a + b) \tag{26}$$

and this holds for power products regardless of the values of a and b. Hence

$$K_{2r} = \frac{(2r)}{n} + \frac{1}{2} \sum_{i=1}^{2r-1} (-1)^i \binom{2r}{i} \frac{(i)(2r - i) - (2r)}{n^{(2)}}$$

$$= \frac{(2r)}{n} \left[1 - \frac{1}{2} \sum_{i=1}^{2r-1} \frac{(-1)^i \binom{2r}{i}}{n - 1} \right] + \frac{1}{2} \sum_{i=1}^{2r-1} (-1)^i \binom{2r}{i} \frac{(i)(2r - i)}{n^{(2)}}.$$

Since

$$\sum_{i=0}^{2r} (-1)^i \binom{2r}{i} = 0 = 2 + \sum_{i=1}^{2r-1} (-1)^i \binom{2r}{i},$$

the coefficient of $\frac{(2r)}{n}$ above is $\frac{n}{n-1}$ and it follows immediately that

$$K_{2r} = \frac{1}{2} \sum_{i=0}^{2r} (-1)^i \binom{2r}{i} \frac{(i)(2r - i)}{n^{(2)}} = \frac{n^2}{n^{(2)}} K_{2r}.$$

This proves the first half of (25) and the second half can be proved in similar fashion, although with considerably more difficulty.

9. **Other Simple Seminvariants which are Invariant under Estimate.** It has been previously remarked (Chapter I, section 2) that the κ system of seminvariants are the seminvariants of minimum degree, those of even weight being of second degree and those of odd weight being of third degree. The κ_{2r}'s are the only seminvariants of degree 2, but for odd weights greater than 7, there exist more than one seminvariant of degree 3. It is not difficult to show that these additional minimum degree seminvariants are also invariant under estimate. The type of proof used could have been applied equally well to obtain the results of the preceding section and indicates that the property of invariance under estimate which is possessed by the κ's is a direct result of their minimum degree property.

Consider the estimate in power product form of any seminvariant of degree 3 and odd weight. Power products of 1, 2 and 3 parts will appear. By the Binet-Waring identities each three part power product (abc) yields a third degree power sum product $(a)(b)(c)$ plus other products of lower degree. Since $(a)(b)(c)$ comes only from (abc) its coefficient must be identical with that of (abc) and will therefore be a constant divided by $n^{(3)}$. The coefficient of each second degree product of power sums will be a sum of terms, the first of which comes from the corresponding two part power product with a coefficient identical with that of the power product, and the others come from the three part power products. Then the coefficient of a second degree product of power sums must be of the form

$$\frac{c_1}{n^{(2)}} + \frac{c_2 + c_3 + \cdots + c_t}{n^{(3)}} = \frac{c_1 n + c_2'}{n^{(3)}}.$$

Similarly the coefficient of the first degree power sum term will be of the form

$$\frac{d_1 n^2 + d_2 n + d_3}{n^{(3)}}.$$

Since the estimate of a seminvariant is a seminvariant, it follows that $d_3 \equiv 0$. This is true because the coefficient of $\dfrac{(r-1)(1)}{n^2}$ must be the coefficient of $\dfrac{(r)}{n}$ multiplied by $-r$. Furthermore $c_2' = d_2 = 0$ for if the contrary be assumed it is immediately possible to break the composite seminvariant into two simple seminvariants, the first being of degree 3 (the original seminvariant) and the second of degree 2. Since for odd weights no seminvariant of degree 2 exists, it follows that any seminvariant of degree 3 and odd weight is invariant under estimate. It is also apparent that the factor $n^3/n^{(3)}$ must appear in the estimate.

10. **Composite Seminvariants which are Invariant under Estimate.** For each weight $r \geqq 4$ there exists a composite seminvariant which is invariant under estimate. For weights 4 and 5 this seminvariant is easily obtained by use

of Table III. Thus for weight 4, form the seminvariant $\lambda_4 + c_{22}\lambda_2^2$. From the table we find that

$$E^{-1}[\lambda_4 + c_{22}\lambda_2^2] = \frac{n^4}{n^{(4)}} l_4 + \frac{n^3}{n^{(4)}} k_4 + c_{22} \frac{n^4}{n^{(4)}} l_2^2 - c_{22} \frac{n}{(n-2)^{(2)}} k_4$$

$$= \frac{n^4}{n^{(4)}} (l_4 + c_{22} l_2^2) + \frac{n}{n^{(4)}} (n^2 - n^{(2)} c_{22}) k_4.$$

If $c_{22} = n^2/n^{(2)}$ the seminvariant is invariant under estimate. This seminvariant is

$$\psi_4 = \lambda_4 + \frac{n^2}{n^{(2)}} \lambda_2^2. \tag{27}$$

In similar fashion we find for weight 5

$$\psi_5 = \lambda_5 + \frac{5n^2}{n^{(2)}} \lambda_3 \lambda_2. \tag{28}$$

For weights > 5 considerably more difficulty is encountered. For weight 6, for example, we consider the seminvariant

$$\lambda_6 + c_{42}\lambda_4\lambda_2 + c_{33}\lambda_3^2 + c_{222}\lambda_2^3.$$

By use of table III we obtain

$$E^{-1}[\lambda_6 + c_{42}\lambda_4\lambda_2 + c_{33}\lambda_3^2 + c_{222}\lambda_2^3] = \frac{n^6}{n^{(6)}} (l_6 + c_{42} l_4 l_2 + c_{33} l_3^2 + c_{222} l_2^3) + \Phi,$$

where Φ is a sum of other seminvariants with coefficients which are functions of n and c_{42}, c_{33}, c_{222}. Now there are only four linearly independent seminvariants of weight 6 and it is necessary that one of these involve the term $(1)^6/n^6$. By an argument analogous to that of the previous section this term cannot appear in Φ and therefore Φ is expressible in terms of three or fewer seminvariants. Actually three are necessary and equating the coefficients of these to zero the values of c_{42}, c_{33} and c_{222} are uniquely determined. The result is somewhat lengthy and scarcely of sufficient interest to record here.

The same sort of procedure can be used for determining seminvariants of higher order which are invariant under estimate, but the labor of computation becomes very great.

It is possible to obtain moment functions which are invariant under estimate by means of a set of equations given by Dwyer [13; 38–39]. These equations connect the coefficients of a general isobaric moment function and the coefficients of the expected value of that function. In his notation if, for example,

$$f_4 = a_4(4) + 4a_{31}(3)(1) + 3a_{22}(2)^2 + 6a_{211}(2)(1)^2 + a_{1111}(1)^4,$$

then

$$E[f_4] = b_4 n\mu_4' + 4b_{31} n^{(2)} \mu_3' \mu_1' + 3b_{22} n^{(2)} \mu_2'^2 + 6b_{211} n^{(3)} \mu_2' \mu_1'^2 + n^{(4)} b_{1111} \mu_1'^4,$$

wherein:

$$
\begin{aligned}
a_4 + 4a_{31} + 3a_{22} + 6a_{211} + a_{1111} &= b_4 \quad , \\
a_{31} \qquad + 3a_{211} + a_{1111} &= b_{31} \quad , \\
a_{22} + 2a_{211} + a_{1111} &= b_{22} \quad , \\
a_{211} + a_{1111} &= b_{211} \quad , \\
a_{1111} &= b_{1111} \, .
\end{aligned} \tag{29}
$$

The problem at hand demands that

$$
E\left[na_4 \frac{(4)}{n} + 4n^2 a_{31} \frac{(3)(1)}{n^2} + 3n^2 a_{22} \frac{(2)^2}{n^2} + 6n^3 a_{211} \frac{(2)(1)^2}{n^3} + n^4 a_{1111} \frac{(1)^4}{n^4} \right]
$$

$$
= \lambda [na_4 \mu_4' + 4n^{(2)} a_{31} \mu_3' \mu_1' + 3n^{(2)} a_{22} \mu_2'^2 + 6n^{(3)} a_{211} \mu_2' \mu_1'^2 + n^{(4)} a_{1111} \mu_1'^4]
$$

so that the equations (29) become

$$
\begin{aligned}
n^4 a_{1111} &= \lambda n^{(4)} a_{1111} \, , \\
n^3 a_{211} &= \lambda n^{(3)} (a_{211} + a_{1111}), \\
n^2 a_{22} &= \lambda n^{(2)} (a_{22} + 2a_{211} + a_{1111}), \\
n^2 a_{31} &= \lambda n^{(2)} (a_{31} + 3a_{211} + a_{1111}), \\
na_4 &= \lambda n (a_4 + 4a_{31} + 3a_{22} + 6a_{211} + a_{1111}),
\end{aligned}
$$

and from these equations a_4, a_{31}, a_{22}, a_{211} can be found in terms of a_{1111}. Obviously there is only one solution if none of the a's are zero. In general, for any weight r, a similar system of equations can be found and they determine the coefficients of a moment function of weight r which is invariant under estimate. It appears that this moment function is always a seminvariant although no proof of the fact has been found. The moment functions of weight 4, 5 and 6 obtained by this method are identical with ψ_4, ψ_5 and ψ_6 defined above.

Conclusion. The results of this paper include:

1. A demonstration of the fact that the theory of statistical seminvariants is identical with the theory of algebraic seminvariants.
2. The introduction of new statistical seminvariants.
3. Simplification of the computation of estimates.
4. Proof that the estimate of any seminvariant is also a seminvariant.
5. Proof of the existence of a trio of seminvariants with the same numerical coefficients.
6. A discussion of seminvariants which are invariant under estimate.

Many thanks are due Professor P. S. Dwyer for his able guidance in the preparation of this paper and to Professors C. C. Craig and J. A. Nyswander for helpful comments.

REFERENCES

[1] L. E. DICKSON, *Algebraic Invariants* (1914).

[2] E. B. ELLIOTT, *An Introduction to the Algebra of Quantics* (1895). Reference is to the second edition (1913).

[3] J. HAMMOND, "On the Solution of the Differential Equation of Sources," *Amer. Jour. of Math.*, Vol. 5 (1882), pp. 218–227.

[4] P. A. MACMAHON, "Seminvariants and Symmetric Functions," *Amer. Jour. of Math.*, Vol. 6 (1884), pp. 131–163.

[5] BURNSIDE and PANTON, *Theory of Equations* (1881). Reference is to the 1904 edition, Vol. II.

[6] P. S. DWYER, "Combined Expansions of Products of Symmetric Power Sums and Sums of Symmetric Power Products with Applications to Sampling." Part I. *Annals of Math. Stat.*, Vol. IX, 1, (1938), pp. 1–47. Part II, Vol. IX, 2, (1938), pp. 97–132.

[7] T. N. THIELE, *Theory of Observations* (1903).

[8] N. P. BERTILSEN, "On the Compatibility of Frequency Constants and on Presumptive Laws of Errors." *Skandinavisk Aktuarietidskrift*, Vol. 10 (1927), pp. 129–156.

[9] A. CAYLEY, "A Memoir on Seminvariants." *Amer. Jour. of Math.* Vol. 7 (1885), pp. 1–25.

[10] T. N. THIELE, *Almindelig Iagttagelseslaere* (1889).

[11] A. A. TSCHUPROW, *Grundbegriffe und Grundprobleme der Korrelationstheorie* (1925).

[12] R. A. FISHER, "Moments and Product Moments of Sampling Distributions." *Proc. London Math. Soc.*, Vol. 2 (30), (1929), pp. 199–238.

[13] P. S. DWYER, "Moments of Any Rational Integral Isobaric Sample Moment Function." *Annals of Math. Stat.*, Vol. 8 (1937), pp. 21–65.

[14] J. R. ROE, "Interfunctional Expressibility Tables of Symmetric Functions." Distributed by Syracuse University (1931).

[15] FAÁ DE BRUNO, *Theorie des Formes Binaires* (1876).

UNIVERSITY OF MICHIGAN,
ANN ARBOR, MICHIGAN.

THE ERRORS INVOLVED IN EVALUATING CORRELATION DETERMINANTS

By Paul G. Hoel

1. **Introduction.** Many statistical problems require for their solution the evaluation of correlation determinants. The method usually employed for such evaluation is that of Chio,[1] in which the order of the determinant is reduced by successive operations with selected pivotal elements. The repeated multiplications and subtractions involved in the method necessitate rounding off the elements in the successively reduced determinants. The calculated value of the original determinant is therefore in error; and so the question naturally arises as to the magnitude of this error.

Previous attempts to answer this question seem to be satisfied with finding an upper bound for the magnitude of the difference between the value of the original determinant and its value after its elements have been rounded off. Moreover, this bound is expressed in terms of the errors in the elements and the minors of the original determinant, whose values are assumed to be known exactly from calculation. However, several reductions are often needed before the value of the determinant can be obtained; and furthermore the minors are subject to the same type of errors as the determinant itself. The problem, therefore, is to find an upper bound for the magnitude of the difference between the final calculated value of the determinant and the determinant itself which involves only calculated quantities.

This paper treats the problem from two different points of view. In the first part an upper bound is obtained for the magnitude of the error. In the second part the first order error terms are given more detailed consideration, with the result that an upper probability bound is obtained for the error.

2. **Absolute Bounds.** Consider the correlation determinant $\Delta = | r_{ij} |$. To evaluate Δ by the method of Chio, it is convenient to select diagonal elements as pivots. It will be assumed without loss of generality that the upper left diagonal element is always chosen as the pivotal element in each reduction. After each reduction, elements are rounded off to a fixed decimal accuracy. Let a_{ij}^k represent the element i,j after the k-th reduction, x_{ij}^k the difference between the rounded value of element a_{ij}^k and a_{ij}^k itself. After k reductions, we arrive at the determinant

$$F^k = \begin{vmatrix} a_{k+1\,k+1}^k + x_{k+1\,k+1}^k & & \\ & \ddots & \\ & & a_{nn}^k + x_{nn}^k \end{vmatrix}$$

[1] See for example, Whittaker and Robinson *Calculus of Observations*, p. 71.

By treating F^k as a function of the x^k, it may be expanded by Taylor's formula as follows:

$$(1) \qquad F^k = A^k + \sum_{i,j=k+1}^{n} x_{ij}^k A_{ij}^k + \frac{1}{2!} \sum_{k+1}^{n}\sum x_{ij}^k x_{pq}^k A_{ijpq}^k + \cdots,$$

where A^k is the value of F^k for all x^k zero, A_{ij}^k is the cofactor of a_{ij}^k in A^k, etc.

For a determinant of order n, the value of the determinant obtained after a single reduction is the value of the original determinant multiplied by the $n - 2$ power of the pivotal element used. Applying this to F^k, it follows that

$$\begin{aligned} A^k &= (a_{kk}^{k-1} + x_{kk}^{k-1})^{n-k-1} F^{k-1} = H_k^{n-k-1} F^{k-1} \\ A_{ij}^k &= H_k^{n-k-2} F_{ij}^{k-1} \\ A_{ijpq}^k &= H_k^{n-k-3} F_{ijpq}^{k-1}, \end{aligned}$$

etc., where the exponents of H_k are ordinary exponents rather than notation. Substituting in (1),

$$F^k = H_k^{n-k-1} F^{k-1} + H_k^{n-k-2} \sum_{k+1}^{n} x_{ij}^k F_{ij}^{k-1} + \frac{1}{2!} H_k^{n-k-3} \sum_{k+1}^{n}\sum x_{ij}^k x_{pq}^k F_{ijpq}^{k-1} + \cdots$$

In order to express F^k in terms of the original determinant, this expansion will be condensed by means of the following operational notation.

$$(2) \qquad F^k = (1 + D + D^2 + \cdots + D^{n-k}) H_k^{n-k-1} F^{k-1},$$

where D^i operates on $H_k^{n-k-1}F^{k-1}$ by reducing the exponent of H_k^{n-k-1} by i units, by summing from $k + 1$ to n the product of i terms in x^k with the corresponding cofactors of F^{k-1}, and dividing the result by factorial i. Using this as a recursion formula,

$$F^k = (1 + D + \cdots + D^{n-k}) H_k^{n-k-1} (1 + \cdots + D^{n-k+1}) H_{k-1}^{n-k} \cdots (1 + \cdots + D^{n-1}) H_1^{n-2} F^0.$$

However,

$$F^0 = \begin{vmatrix} a_{11} + x_{11} & & \\ & \ddots & \\ & & a_{nn} + x_{nn} \end{vmatrix} = \Delta,$$

since we assume that $x_{ij} = 0$ for our original determinant. Consequently,

$$(3) \qquad F^k = (1 + \cdots + D^{n-k}) H_k^{n-k-1} (1 + \cdots + D^{n-k+1}) H_{k-1}^{n-k} \cdots (1 + \cdots + D^{n-1}) H_1^{n-2} \Delta.$$

Since D^i operates on F^{k-1} in (2) to extract the proper cofactor of i less rows than in F^{k-1}, which in turn reduces the exponent of all factors H_{k-1} in the expansion of F^{k-1} by i units, D^i reduces the exponent of all H's following it in the expansion of F^k in (3) by i units.

Following these rules of operation, and expanding so as to collect terms of the same degree in the x's, we may write

$$
(4) \qquad \begin{aligned} F^k = H_k^{n-k-1} \cdots H_1^{n-2}\Delta + H_k^{n-k-2} \cdots H_1^{n-3} \text{ (terms in } x_{ij}) + \\ H_k^{n-k-3} \cdots H_1^{n-4} \text{ (terms in } x_{ij}x_{pq}) + \cdots . \end{aligned}
$$

Letting $H = H_k H_{k-1} \cdots H_1$ and $C = H_k^{n-k-1} \cdots H_1^{n-2}$, we may write

$$
I = F^k - C\Delta = C\left[\frac{1}{H}(\text{terms in } x_{ij}) + \frac{1}{H^2}(\text{terms in } x_{ij}x_{pq}) + \cdots\right];
$$

and hence

$$
(5) \qquad J = \frac{F^k}{C} - \Delta = \frac{1}{H}(\text{terms in } x_{ij}) + \frac{1}{H^2}(\text{terms in } x_{ij}x_{pq}) + \cdots .
$$

Now J is the difference between the calculated value of Δ, using Chio's reduction method and rounding off after each reduction, and the true value of Δ. We are interested in finding an upper bound for the magnitude of J. To accomplish this we shall first overestimate the number of terms in the various sums of (5), then find an upper bound for the magnitude of the terms in these sums, and finally combine the two results.

In counting terms by means of (3), we may ignore the H's since they merely serve as coefficients of the x's. Therefore consider the nature of the terms in

$$
(1 + \cdots + D^{n-k})(1 + \cdots + D^{n-k+1}) \cdots (1 + \cdots + D^{n-1})\Delta.
$$

Now $(1 + \cdots + D^s)\Delta$ contains the sums $\sum_{n-s+1}^{n} x_{ij}\Delta_{ij}$, $\frac{1}{2!}\sum\sum_{n-s+1}^{n} x_{ij}x_{pq}\,\Delta_{ijpq}$, etc.; hence it contains s^2 terms in x_{ij}, $\frac{s^2(s-1)^2}{2}$ terms in $x_{ij}x_{pq}$, etc. Each of these is not greater than s^2, ${}_{s^2}C_2$, etc.; consequently, the number of terms of each type is not greater than the coefficient of the corresponding power of D in the expansion of $(1 + D)^{s^2}$. Therefore,

$$
(6) \qquad (1 + D)^{(n-k)^2}(1 + D)^{(n-k+1)^2} \cdots (1 + D)^{(n-1)^2} = (1 + D)^m,
$$

where $m = (n - k)^2 + \cdots + (n - 1)^2$, contains at least as many terms of each type as are found in the expansion of F^k. This gives us the desired overestimate of the number of terms in the various sums of (5).

In finding upper bounds for the magnitudes of terms, it is to be noted that (4) is written with all common factors extracted from each set of terms of the same degree in the x's. In the parenthesis containing terms consisting of the product of r x's, the first sum will have unity for its coefficient while the last sum will have $H_k^r H_{k-1}^r \cdots H_2^r$ as coefficient, with all sums between having as coefficients products of H's with exponents $\leq r$. Hence an upper bound for all coefficients in this parenthesis may be written as $\bar{H}^r$, where $\bar{H}$ is the magnitude of the product of those H's whose magnitude is greater than unity, but unity if none exceeds

unity. Now terms in x_{ij} are multiplied by Δ_{ij}, those in $x_{ij}x_{pq}$ by Δ_{ijpq}, etc.; therefore let $\bar{\Delta}_{ij}$, $\bar{\Delta}_{ijpq}$, etc., be the absolute values of the largest in magnitude of such cofactors. With this notation for upper bounds for magnitudes of terms, and (6) giving an upper bound for the number of terms, we may write an upper bound for the magnitude of J as follows:

$$|J| \leq \left(\frac{\bar{H}}{H}\epsilon\right) {}_mC_1 \bar{\Delta}_{ij} + \left(\frac{\bar{H}}{H}\epsilon\right)^2 {}_mC_2 \bar{\Delta}_{ijpq} + \cdots, \tag{7}$$

where $\epsilon \geq |x|$ is the maximum error of rounding. This result is valid for any determinant with real elements. All quantities on the right are available from calculations except the $\bar{\Delta}$; consequently this upper bound will be useful only if satisfactory bounds exist for the minors of the determinant. It can be shown that (7) holds for any minor of Δ, say Δ_{uv}, if the $\bar{\Delta}$ have uv added as subscripts; and therefore it may be applied to the question of the accuracy of least square solutions.

For the correlation determinant Δ it can be shown that the magnitude of a minor of order $n - k$ is bounded by $k!/2^{\frac{1}{2}k}$ for k even and $k!/2^{\frac{1}{2}(k-1)}$ for k odd. Setting $a = \dfrac{\bar{H}}{H}\epsilon$ and substituting these bounds in (7),

$$\begin{aligned} |J| &\leq am + a^2 {}_mC_2 \frac{2!}{2} + a^3 {}_mC_3 \frac{3!}{2} + a^4 {}_mC_4 \frac{4!}{2^2} + \cdots \\ &\leq am + \frac{a^2m^2}{2} + \frac{a^3m^3}{2} + \frac{a^4m^4}{2^2} + \cdots \\ &\leq am + \frac{a^2m^2}{2} + \frac{a^3m^3}{2(1 - am)}, \end{aligned} \tag{8}$$

for $am < 1$. Since am is obtainable from the calculations for Δ, this is the desired upper bound for the error in question.

3. **Probability Bounds.** In order to find probability bounds for this error, it will be necessary to expand the H's since they involve the variables x. Consider $H_k = a_{kk}^{k-1} + x_{kk}^{k-1}$. Since a_{kk}^{k-1} came from repeated reductions of Δ, it is expressible in terms of the x's and the minors of Δ. To obtain this expansion of H_k consider

$$G^s = \begin{vmatrix} a_{k-s+1\,k-s+1}^{k-s} + x_{k-s+1\,k-s+1}^{k-s} & & \\ & \ddots & \\ & & a_{kk}^{k-s} + x_{kk}^{k-s} \end{vmatrix}$$

Using the same methods as for F^k, this may be written as

$$G^s = B^s + \sum_{k-s+1}^{k} x_{ij}^{k-s} B_{ij}^s + \frac{1}{2!} \sum\sum_{k-s+1}^{k} x_{ij}^{k-s} x_{pq}^{k-s} B_{ijpq}^s + \cdots,$$

where B^s is the value of G^s for all x^{k-s} zero, etc., and where $B^s = H_{k-s}^{s-1}G^{s+1}$, $B_{ij}^s = H_{k-s}^{s-2}G_{ij}^{s+1}$, etc. Substituting,

$$G^s = H_{k-s}^{s-1}G^{s+1} + H_{k-s}^{s-2}\sum x_{ij}^{k-s}G_{ij}^{s+1} + \frac{1}{2!}H_{k-s}^{s-3}\sum\sum x_{ij}^{k-s}x_{pq}^{k-s}G_{ijpq}^{s+1} + \cdots$$

Using operational notation here also, this may be written as

$$G^s = (1 + E + E^2 + \cdots + E^s)H_{k-s}^{s-1}G^{s+1},$$

where the E's operate the same as the D's, except that sums are taken from $k - s + 1$ to k rather than from $n - s + 1$ to n. Treating this as a recursion formula,

$$H_k = G^1 = (1 + E)H_{k-1}^0(1 + E + E^2)H_{k-2}^1 \cdots (1 + \cdots + E^{k-1})H_1^{k-2}G^k.$$

However,

$$G^k = \begin{vmatrix} a_{11} + x_{11} & & \\ & \ddots & \\ & & a_{kk} + x_{kk} \end{vmatrix} = \begin{vmatrix} a_{11} & & \\ & \ddots & \\ & & a_{kk} \end{vmatrix} = \Delta_k.$$

Consequently,

$$H_k = (1 + E)H_{k-1}^0(1 + E + E^2)H_{k-2}^1 \cdots (1 + \cdots + E^{k-1})H_1^{k-2}\,\Delta_k. \tag{9}$$

Since the E's operate on the following H's to reduce their exponents, the number of terms of various types, that is, of various degrees in the x's, will not be decreased if the order of H's is disregarded and their exponents held fixed. Therefore consider

$$H_k' = (1 + E)(1 + E + E^2) \cdots (1 + \cdots + E^{k-1})\Delta_k H_{k-1}^0 \cdots H_1^{k-2} \tag{10}$$

as an ordinary recursion formula in the H's for overestimating the number of terms of various types. If (10) is substituted for successive H's within itself in a systematic manner until no H's remain, it will be found that

$$\begin{aligned} H_k' = {} & (1 + E) \cdots (1 + \cdots + E^{k-1})\Delta_k \\ & [(1 + E) \cdots (1 + \cdots + E^{k-3})\Delta_{k-2}]^{2^0} \cdots [(1 + E)\Delta_2]^{2^{k-4}}[\Delta_1]^{2^{k-3}}. \end{aligned} \tag{11}$$

To merely count terms it is permissible to combine like terms to give

$$\begin{aligned} H_k' &= (1 + E)^{2+2+2^2+\cdots+2^{k-4}}(1 + E + E^2)^{2+2+2^2+\cdots+2^{k-5}} \cdots (1 + \cdots + E^{k-1})K \\ &= (1 + E)^{2^{k-3}}(1 + E + E^2)^{2^{k-4}} \cdots (1 + \cdots + E^{k-1})K, \end{aligned}$$

where K is the product of the Δ's. Since the E's operate like the D's, the same arguments as those used to arrive at (6) may be used to replace $(1 + E + \cdots + E^s)$ by $(1 + E)^{s^2}$ for overestimating the number of terms. Hence, the number of terms of various types in H_k is not greater than those in

$$(1 + E)^{2^{k-3}}(1 + E)^{2^2 \cdot 2^{k-4}} \cdots (1 + E)^{(k-2)^2 \cdot 2^0}(1 + E)^{(k-1)^2} = (1 + E)^{w_k},$$

where $w_k = 2^{k-3} + 2^2 \cdot 2^{k-4} + \cdots + (k-2)^2 \cdot 2^0 + (k-1)^2$. Therefore the number of terms of various types in $H_k^{n-k-1} \cdots H_1^{n-2}$ is not greater than in

$$(12) \qquad (1+E)^{(n-k-1)w_k+(n-k)w_{k-1}+\cdots+(n-2)w_1} = (1+E)^t.$$

It is easily shown that t can be condensed into the form

$$(13)\ t = [2^{k-2}(n-k)-1] + 2^2[2^{k-3}(n-k)-1] + \cdots + (k-1)^2[2^0(n-k)-1].$$

From (3) it is evident that the number of terms of various types in F^k will not be greater than those in the expansion of F^k when the exponents of the H's are held fixed. But from (6) we have an upper bound for the number of terms arising from the D's, and from (12) those arising from the H's; hence the number of terms in question will certainly be bounded by those in

$$(14) \qquad (1+D)^{m+t} = (1+D)^{\mu}.$$

Now consider the magnitude of terms. The terms arising from the operation of D's contain minors of Δ as factors, while those arising from the operation of E's contain minors of Δ_i, where i ranges from 1 to k. Let Δ'_{ij}, etc., denote an upper bound for the magnitudes of all such minors of the same number of subscripts. It is easily shown that Δ' with $2r$ subscripts is not less than the magnitude of the product of several minors whose subscripts total $2r$ in number. The terms of various types also contain as factors products of the constant terms in the H's. The constant term in H_k, which will be denoted by h_k, can be obtained from (11) by operating with all ones since it will be unaffected by disregarding the order of operation. Hence,

$$h_k = \Delta_k \Delta_{k-2} \Delta_{k-3}^2 \cdots \Delta_2^{2^{k-4}} \Delta_1^{2^{k-3}}.$$

Since the Δ_i are principal minors of a positive definite determinant with no element greater than unity, h_k has unity for an upper bound. Thus, an upper bound for the magnitude of any term in the product of i x's will be ϵ^i times Δ' with $2i$ subscripts.

With upper bounds now available for the number of terms and the magnitudes of terms, we are in a position to consider the complete expansion of I in which the coefficients of the x's will be constants rather than H's. Evidently the terms in x_{ij} will come from the terms in x_{ij} of (4) with the H's replaced by the constant terms in their expansions. If Z denotes these terms, then

$$(15) \qquad \begin{aligned} Z = h_k^{n-k-2} \cdots h_1^{n-3}\Bigg[\sum_{k+1}^{n} x_{ij}^{k}\Delta_{ij} + h_k \sum_{k}^{n} x_{ij}^{k-1}\Delta_{ij} \\ + \cdots + h_k \cdots h_2 \sum_{2}^{n} x_{ij}^{1}\Delta_{ij}\Bigg]. \end{aligned}$$

Now consider an upper bound for $|I - Z|$. Since $I - Z$ involves only terms in the product of two or more x's, we need consider an upper bound for such terms only. From the results of the two preceding paragraphs, we obtain

$$|I - Z| \leq \epsilon^2{}_{\mu}C_2\Delta'_{ijpq} + \epsilon^3{}_{\mu}C_3\Delta'_{ijpquv} + \cdots.$$

But from the paragraph containing (8), bounds are available for the Δ'; hence

$$|I - Z| \leq \epsilon^2 {}_\mu C_2 + \epsilon^3 {}_\mu C_3 \frac{3!}{2} + \epsilon^4 {}_\mu C_4 \frac{4!}{2^2} + \cdots$$
$$\leq \frac{\epsilon^2 \mu^2}{2} + \frac{\epsilon^3 \mu^3}{2(1 - \epsilon\mu)} = \Phi,$$

for $\epsilon\mu < 1$. Since Z is of order ϵ, Φ will ordinarily be small compared with Z; therefore consider the nature of the distribution of Z.

If we write $Z = a_1 x_1 + \cdots + a_p x_p$, then, since the x's are independently distributed with rectangular distributions, it is easily shown that $\mu_2 = \frac{\epsilon^2}{3} \sum a_i^2$, $\alpha_3 = 0$, $\alpha_4 = 3 - \frac{6}{5} \sum a_i^4 / (\sum a_i^2)^2$. If the a_i are approximately equal in magnitude, then α_4 is approximately equal to $3 - 1/p$. But from (15) $p \geq \frac{1}{2}(n - k)^2 + \cdots + \frac{1}{2}(n - 1)^2$, which is sufficiently large for determinants employing Chio's method to justify the assumption that Z is approximately normally distributed. Setting $L = h_k^{n-k-2} \cdots h_1^{n-3}$,

$$\mu_2 = \frac{L^2 \epsilon^2}{3} \left[\left(\sum_{k+1}^{n} \Delta_{ii}^2 + 4 \sum_{i<j} \Delta_{ij}^2 \right) + \cdots + h_k^2 \cdots h_2^2 \left(\sum_{2}^{n} \Delta_{ii}^2 + 4 \sum_{i<j} \Delta_{ij}^2 \right) \right]$$
$$\leq \frac{2\epsilon^2}{3} [(n - k)^2 + \cdots + (n - 1)^2 - \tfrac{1}{2}\{(n - k) + \cdots + (n - 1)^2\}]$$
$$\leq \frac{2\epsilon^2}{3} \left[(n - k)^2 + \cdots + (n - 1)^2 - \frac{k}{4}(2n - k - 1) \right] = \Psi^2.$$

Hence, the probability is $> .95$ that $|Z| < 2\Psi$. Since $|I - Z| \leq \Phi$, the probability is $> .95$ that $|I| < 2\Psi + \Phi$; and therefore the probability is $> .95$ that

$$|J| < \frac{2\Psi + \Phi}{C}. \tag{16}$$

This inequality will usually give a smaller bound for $|J|$ than (8). However, when Δ is small the H's may be small, with the result that C will be small and (16) may not give a satisfactory bound for $|J|$. In such cases the bound given by (8) may not prove satisfactory either.

4. **Example.** Consider a correlation determinant of order 7 in which the elements are accurate to 4 decimal places. If Chio's reduction method is applied until a 2 rowed determinant is obtained, then $n = 7$, $k = 5$, $\epsilon = .00005$, $m = 90$, $\mu = 176$, $\Psi = .00005\sqrt{160}/3$, and we obtain from (8) that

$$|J| < \left(\frac{\bar{H}}{H}\right) .0045 + \left(\frac{\bar{H}}{H}\right)^2 .00001 + \left(\frac{\bar{H}}{H}\right)^3 \frac{.00000005}{1 - .0045\,\bar{H}/H}$$

where $\bar{H}/H$ is obtained from calculations involved in evaluating the determinant. From (16) we obtain that the probability is $>.95$ that

$$|J| < \frac{.0008}{C}.$$

The relative advantage of the second inequality over the first depends on the size of the pivotal elements, as does the usefulness of either inequality.

UNIVERSITY OF CALIFORNIA AT LOS ANGELES

THE CUMULATIVE NUMBERS AND THEIR POLYNOMIALS

By P. S. Dwyer

In a recent paper [1] the author has shown how the moments of a distribution can be obtained from the last entries of cumulative columns with the use of multiplication by certain numbers. These numbers may be called "cumulative numbers." It is the aim of this paper to show how these numbers can be obtained from the expansion of x^s in terms of factorials of the s-th order and to demonstrate properties of the polynomials of which these numbers are the coefficients.

TABLE 1

Successive Frequency Cumulations

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
X	x	f_x	C^1	C^2	C^3	C^4	C^5
$a + 6$	6	64	64	64	64	64	64
$a + 5$	5	192	256	320	384	448	512
$a + 4$	4	240	496	816	1200	1648	2160
$a + 3$	3	160	656	1472	2672	4320	6480
$a + 2$	2	60	716	2188	4860	9180	15660
$a + 1$	1	12	728	2916	7776	16956	32616
a	0	1	729	3645	11421	28377	60993

1. **The values $C_i^j(u_x)$.** We use the notation $C_i^j(u_x)$ of the previous paper [1,289] to express the columnar cumulated entries. The j indicates the order of the cumulation while the i indicates the number of the term, counting from the bottom of the column. Thus in Table I, which presents the cumulations of a frequency distribution used in the previous paper [1,289], $C_1^1 = 729$; $C_1^2 = 3645$; $C_2^2 = 2916$; $\cdots$, $C_4^5 = 6480$, etc. Now if $k + 1$ values of x are spaced at unit distances and if the smallest value of x is 0, it can be shown that

$$C_1^1 = \sum_0^k u_x; \quad C_1^2 = \sum_0^k (x + 1)u_x; \quad C_2^2 = \sum_0^k xu_x; \quad C_1^3 = \sum_0^k \frac{(x + 2)(x + 1)}{2!} u_x;$$

$$C_2^3 = \sum_0^k \frac{(x + 1)x}{2!} u_x; \qquad C_3^3 = \sum_0^k \frac{x(x - 1)}{2!} u_x$$

and, in general, $j > 0$ and $j + 1 \geq i$,

$$C_i^{j+1} = \sum_{x=0}^k \frac{(x + j + 1 - i)^{(j)}}{j!} u_x. \tag{1}$$

Similarly if k values of x are spaced at unit distances and if the smallest value of x is 1, it can be shown that

$$C_1^1 = \sum_1^k u_x; \quad C_1^2 = \sum_1^k xu_x; \quad C_2^2 = \sum_1^k (x-1)u_x; \quad C_1^3 = \sum_1^k \frac{(x+1)x}{2!} u_x;$$

$$C_2^3 = \sum_1^k \frac{x(x-1)}{2!} u_x; \qquad C_3^3 = \sum_1^k \frac{(x-1)(x-2)}{2!} u_x$$

and, in general, $j > 0$ and $j + 1 \geq i$,

$$(2) \qquad C_i^{j+1} = \sum_{x=1}^k \frac{(x+j-i)^{(j)}}{j!} u_x.$$

It is to be noted that the coefficients of u_x in (2) could be obtained from the coefficients of u_x in (1) by the substitution $x + 1 = x'$.

2. **The powers in terms of factorials of the s-th order.** If the s-th powers can be expressed in terms of factorials of the s-th order (factorials having s factors) then the moments can be expressed in terms of the cumulations. For example

$$x^2 = \frac{(x+1)x + x(x-1)}{2} \text{ so, from (1)}$$

$$\sum_0^k x^2 f_x = \sum_0^k \frac{(x+1)^{(2)}}{2!} f_x + \sum_0^k \frac{x^{(2)}}{2!} f_x = C_2^3 + C_3^3.$$

And since

$$x^3 = \frac{(x+2)^{(3)} + 4(x+1)^{(3)} + x^{(3)}}{3!}, \text{ we have}$$

$$\sum_0^k x^3 f_x = \sum_0^k \frac{(x+2)^{(3)}}{3!} f_x + 4 \sum_0^k \frac{(x+1)^{(3)}}{3!} f_x + \sum_0^k \frac{x^{(3)}}{3!} f_x = C_2^4 + 4C_3^4 + C_4^4.$$

In general if

$$(3) \quad x^s = \frac{A_{s1}(x+s-1)^{(s)} + A_{s2}(x+s-2)^{(s)} + \cdots + A_{sj}(x+s-j)^{(s)} + \cdots + A_{ss}x^{(s)}}{s!},$$

then

$$(4) \quad \sum_0^k x^s f_x = A_{s1}C_2^{s+1} + A_{s2}C_3^{s+1} + \cdots + A_{sj}C_{j+1}^{s+1} + \cdots + A_{ss}C_{s+1}^{s+1},$$

while if the smallest value of x is 1, we have

$$(5) \quad \sum_1^k x^s f_x = A_{s1}C_1^{s+1} + A_{s2}C_2^{s+1} + \cdots + A_{sj}C_j^{s+1} + \cdots + A_{ss}C_s^{s+1}.$$

These quantities, A_{sj}, in (4) and (5) are simply the coefficients of certain factorials of the s-th order in the expansion of $x^s s!$.

These numbers, for small values of s, are easily obtained. It is possible to use the table and a recursion formula of a previous paper [1,294–295] for larger values of s. It is also possible to obtain these values, without involving cumulative theory, from (3) above.

While doing this we make a more general approach by expanding $(a + x)^s$ in terms of these same factorials with the coefficients now functions of a. This is possible if we add an additional term, $A_{s0}(x + s)^{(s)}$, to the numerator of the right hand side of (3). We have then

$$(6) \quad (a + x)^s = \frac{\begin{array}{r} A_{s0}(x + s)^{(s)} + A_{s1}(x + s - 1)^{(s)} \\ + \cdots + A_{sj}(x + s - j)^{(s)} + \cdots + A_{ss}x^{(s)} \end{array}}{s!}.$$

The determination of the values A_{sj} can be accomplished by purely algebraic means by successive substitution of $x = 0, 1, 2, \cdots s$. In this way we obtain $s + 1$ equations in $s + 1$ unknowns. For example when $s = 2$

$$(a + x)^2 = \frac{A_{20}(x + 2)^{(2)} + A_{21}(x + 1)^{(2)} + A_{22}x^{(2)}}{2!}$$

so that when $x = 0, 1, 2$, we have

$$a^2 = A_{20}\ ;\ (a + 1)^2 = 3A_{20} + A_{21}\ ;\ (a + 2)^2 = 6A_{20} + 3A_{21} + A_{22}\,.$$

The solution is $A_{20} = a^2$; $A_{21} = 2ab + 1$; $A_{22} = b^2$ where $b = 1 - a$. It follows that

$$(a + x)^2 = a^2\frac{(x + 2)^{(2)}}{2!} + (2ab + 1)\frac{(x + 1)^{(2)}}{2!} + b^2\frac{x^{(2)}}{2!} \text{ and hence that}$$

$$\sum_0^k (a + x)^2 f_x = a^2C_1^3 + (2ab + 1)C_2^3 + b^2C_3^3,$$

as indicated in the previous paper [1,293].

When $a = 0$, then $b = 1$ and we have

$\Sigma x^2 f_x = C_2^3 + C_3^3$ while when $a = 1$, $b = 0$ and the right hand side becomes $C_1^3 + C_2^3$.

It follows that the general cumulative numbers might also be defined as the solutions of the $s + 1$ equations in the $s + 1$ unknowns obtained by placing $x = 0, 1, 2, \cdots, s$ in (6).

3. The evaluation of the cumulative numbers. Formal algebraic methods of evaluating equations (6) are somewhat tedious so we use finite difference theory to aid in finding the solution. As in the previous paper [1] we use the notation

$\nabla v_x = v_x - v_{x-1}$ and $\underset{\cdot}{v}_x = \begin{cases} v_x \text{ when } a \le x \le a + k \\ 0 \text{ otherwise} \end{cases}$. We then write, from (6)

$$
(7)\quad \begin{aligned} s!\underline{(a+x)^s} &= A_{s0}\underline{(x+s)^{(s)}} + A_{s1}\underline{(x+s-1)^{(s)}} \\ &\quad + \cdots + A_{sj}\underline{(x+s-j)^{(s)}} + \cdots + A_{ss}\underline{x^{(s)}}. \end{aligned}
$$

We note further that $\nabla^{s+1}\underline{(x+r)^{(s)}} = \begin{Bmatrix} s! \text{ when } r = 0 \\ 0 \text{ when } r \neq 0 \end{Bmatrix}$. We have then

$$
(8)\quad \nabla^{s+1}\underline{(a+j)^s} = A_{sj}\,.
$$

It has been shown in the previous paper [1,292] that

$$
(9)\quad \nabla^{s+1}\underline{(a+j)^s} = \sum_{t=0}^{j} (-1)^t \binom{s+1}{t} (a+j-t)^s
$$

and it appears that the cumulative numbers could be defined by (9). A useful recursion formula has been derived from (9)

$$
(10)\quad \nabla^{s+1}\underline{(a+x)^s} = (a+x)\nabla^s\underline{(a+x)^{s-1}} + (s+1-a-x)\nabla^s\underline{(a+x-1)^{s-1}}.
$$

4. **The cumulative polynomials.** We define the cumulative polynomials to be the polynomials obtained by using the cumulative numbers as coefficients. Thus when $a = 0$,

$P_1 = y$; $P_2 = y + y^2$; $P_3 = y + 4y^2 + y^3$; $P_4 = y + 11y^2 + 11y^3 + y^4$; etc.

It is possible to derive a recursion formula for these polynomials. We use (10) with s replaced by $s + 1$ and $a = 0$ and get

$$
(11)\quad P_{s+1} = \Sigma\nabla^{s+2}\underline{(x)^{s+1}}y^x = \Sigma x\nabla^{s+1}\underline{(x)^s}y^x + \Sigma(s+2-x)\nabla^{s+1}\underline{(x-1)^s}y^x,
$$

which becomes, after some manipulation,

$$
(12)\quad P_{s+1} = (1-y)\Sigma x\nabla^{s+1}\underline{(x)^s}y^x + (s+1)yP_s\,.
$$

To illustrate we get P_4 from $P_3 = y + 4y^2 + y^3$. Now $\Sigma x\nabla^4\underline{(x)^3}y^x = y + 8y^2 + 3y^3$ and $P_4 = (1-y)(y + 8y^2 + 3y^3) + 4y(y + 4y^2 + y^3) = y + 11y^2 + 4y^3 + y^4$. The recursion formula (12) can be expressed also in the form of a differential equation, since $P'_s = \dfrac{d}{d_y}(P_s) = \Sigma x\nabla^{s+1}\underline{(x)^s}y^{x-1}$, as

$$
(13)\quad P_{s+1} = y[(1-y)P'_s + (s+1)P_s].
$$

It can be shown more generally that for any a

$P_{a,0} = 1$; $P_{a,1} = a + by$; $P_{a,2} = a^2 + (2ab + 1)y + b^2y^2$, etc. with

$$
(14)\quad P_{a,s+1} = y(1-y)P'_{a,s} + [a(1-y) + (s+1)y]P_{a,s}
$$

as the recursion formula.

5. **The numerator coefficients in successive derivatives of the logistic function.** Lotka has recently exhibited the coefficients of the numerator terms of suc-

cessive derivatives of the logistic function [2, 160]. These appear to be, aside from sign, the same as the cumulative numbers when $a = 0$. It is shown in this section that these numbers are the cumulative numbers. The scheme is generalized to include the numerator coefficients of the derivatives of a more general function involving the parameter a.

Lotka used the function $\Phi_0 = \dfrac{1}{1 + e^{rt}}$ and obtained $\Phi_1 = \dfrac{re^{rt}}{(1 + e^{rt})^2}$, $\Phi_2 = \dfrac{r^2 e^{rt}(1 - e^{rt})}{(1 + e^{rt})^3}$, etc. The numerical coefficients are the same if $r = 1$ so we might as well use $\Phi_0 = \dfrac{1}{1 + e^x}$. A more general function is the two parameter function

$$\Phi_{a,c} = \frac{e^{ax}}{1 + ce^x}. \tag{15}$$

Let successive derivatives with respect to x be indicated by $\Phi_{a,c,1}$; $\Phi_{a,c,2}$; $\Phi_{a,c,3}$; etc. Then

$$\Phi_{a,c,1} = \frac{e^{ax}[a + c(1 - a)e^x]}{(1 + ce^x)^2},$$

$$\Phi_{a,c,2} = \frac{e^{ax}[a^2 + (-2a^2 + 2a + 1)ce^x + (1 - a)^2c^2e^{2x}]}{(1 + ce^x)^3}.$$

In general,

$$\Phi_{a,c,s} = \frac{e^{ax}Q_{a,c,s}}{(1 + ce^x)^{s+1}} = e^{ax}Q_{a,c,s}(1 + ce^x)^{-s-1}$$

so that

$$\Phi_{a,c,s+1} = \frac{e^{ax}\{(1 + ce^x)[aQ_{a,c,s} + Q'_{a,c,s}] - (s + 1)ce^xQ_{a,c,s}\}}{(1 + ce^x)^{s+2}}$$

and

$$Q_{a,c,s+1} = (1 + ce^x)[aQ_{a,c,s} + Q'_{a,c,s}] - (s + 1)ce^xQ_{a,c,s}. \tag{16}$$

The Q functions can be changed to polynomials with the substitution $e^x = y$. Then derivatives are taken with respect to y and

$$P_{a,c,s+1} = (1 + cy)[aP_{a,c,s} + yP'_{a,c,s}] - (s + 1)cyP_{a,c,s}. \tag{17}$$

When $c = -1$, this becomes formula (14) and since $P_{a,0} = 1$, it follows that the numbers of the present section are generalized cumulative numbers. When $c = 1$ and $a = 0$ we have the numbers found by Lotka.

It can be shown, further, that the c coefficient of y^j is c^j. It follows that the absolute values of the coefficients, when $c = 1$ and when $c = -1$, are the same.

6. Formulas for Σx^s. A formula for the sums of the s-th powers of the integers from 1 to k is obtained by summing (3). We get

$$\sum_1^k x^s = A_{s1} \sum_1^k \frac{(x+s-1)^{(s)}}{s!} + \cdots \tag{18}$$
$$+ A_{sj} \sum_1^k \frac{(x+s-j)^{(s)}}{s!} + \cdots + A_{ss} \sum_1^k \frac{x^{(s)}}{s!}$$

from which

$$\sum_1^k x^s = A_{s1} \frac{(k+s-1)^{(s+1)}}{(s+1)!} + \cdots \tag{19}$$
$$+ A_{sj} \frac{(k+s-j)^{(s+1)}}{(s+1)!} + \cdots + A_{ss} \frac{k^{(s+1)}}{(s+1)!},$$

or

$$\sum_1^k x^s = \sum_{j=1}^s A_{sj} \frac{(k+s-j)^{(s+1)}}{(s+1)!} = \frac{1}{(s+1)!} \sum_{j=1}^s (k+s-j)^{(s+1)} \nabla^{s+1} \underline{(j)^s} \tag{20}$$
$$= \sum_{j=1}^s C_j^{s+1}(1) \nabla^{s+1} \underline{(j)^s}.$$

For example

$$\sum_1^k x^2 = \frac{(k+2)^{(3)} + (k+1)^{(3)}}{3!} = \frac{k(k+1)(2k+1)}{6},$$
$$\sum_1^k x^3 = \frac{(k+3)^{(4)} + 4(k+2)^{(4)} + (k+1)^{(4)}}{4!} = \frac{k^2(k+1)^2}{4}.$$

More generally the values of $\sum_a^{a+k} x^s$ can be evaluated by

$$\sum_a^{a+k} x^s = \frac{1}{(s+1)!} \sum_{j=0}^s (k+s-j)^{(s+1)} \nabla^{s+1} \underline{(a+j)^s} = \sum_{j=0}^s C_{j+1}^{s+1}(1) \nabla^{s+1} \underline{(a+j)^s}. \tag{21}$$

7. **Summary.** It is shown how the cumulative numbers and the cumulative polynomials may be obtained in a variety of ways. Of special interest is the fact that the cumulative numbers can be obtained by expanding powers in terms of factorials and hence they might be called factorial coefficients of a kind. It is also possible, though it is not within the scope of this paper, to establish interesting relations between the cumulative numbers and the multinomial coefficients, the usual factorial coefficients, the difference of 0, etc.

REFERENCES

[1]. P. S. Dwyer. "The Computation of Moments with the use of Cumulative Totals." *Annals of Math. Stat.* Vol. IX, no. 4, Dec. 1938, pp. 288–304. A more extensive bibliography is available here.

[2]. A. J. Lotka. "An Integral Equation in Population Analysis." *Annals of Math. Stat.* Vol. X, no. 2, June 1939, pp. 144–161.

University of Michigan
Ann Arbor, Michigan

ENUMERATION AND CONSTRUCTION OF BALANCED INCOMPLETE BLOCK CONFIGURATIONS[1]

By Gertrude M. Cox

1. **Introduction.** One of the general problems of experimental design is to avoid extraneous effects in making desired comparisons. The method employed is to use experimental materials as nearly homogeneous as possible. Such materials, however, are seldom available in large quantities. On the contrary, field soils vary in fertility from block to block, animals vary with both litter and sex, and leaves on one young plant differ from those on another. Differences between blocks, between litters and sex, and between plants, being irrelevant to the comparisons usually contemplated, must be avoided.

When the number of treatments to be compared is small, well known methods of design, such as the Latin square or randomized complete block, are available and efficient. As the number of treatments increases, however, these designs tend to become less efficient through failure to eliminate heterogeneity. Furthermore, they become cumbersome, the Latin square design requiring replicates equal in number to the treatments and the complete block design providing that each treatment occur in every block. (Blocks are defined as an assemblage of experimental units chosen to be as nearly alike as possible.)

Because of such limitations, several modifications of the complete block design have been devised. These new designs all have the common characteristic that the experimental material is divided into groups or blocks containing fewer units than the number of treatments to be compared. These more homogeneous small blocks are referred to as incomplete blocks.

It is desirable to have all comparisons between pairs of treatments made with equal accuracy. This requires of the design that every pair of treatments occur in the same block an equal number of times. Such a design is referred to as balanced. Balanced incomplete block designs can be arranged (for any given number of treatments) only for certain combinations of block size and number of replications.[2]

The construction of balanced incomplete block designs is mathematically a part of the theory of configurations. A configuration is an assemblage of elements into sets, each element occurring in the same number of sets, and each

[1] A revision of an expository paper presented under a different title at a joint meeting of the Institute of Mathematical Statistics and Biometric Section of the American Statistical Association, December 27, 1939.

[2] Numerous additional designs are available in the partially balanced incomplete blocks [3].

set containing the same number of elements. The configurations to be considered here are the complete configurations, i.e., those in which each element occurs an equal number of times in the same set with every other element. It would be useful to know, (a) what configurations (within the useful range) exist. (b) how these configurations may be constructed.

The typical requirement of the experimenter is this: "I wish to test t treatments and can use blocks of size $k(t > k)$. I should like a design which will involve as little experimental material as feasible." The designer must then determine what configuration of t elements in sets of k will satisfy the incidence relation that each pair of elements occur together in a set an equal number of times, and for which the total number of sets is a minimum. There are still many configurations which the experimenter needs but which have not as yet been constructed.

In order better to explain the construction of these balanced incomplete block designs, it is essential to specify the underlying combinatorial problems. A configuration satisfying the condition of balance can be obtained by writing down all possible combinations, b, of the t elements taken k at a time,

$$b = {}_tC_k = \frac{t!}{k!(t-k)!}.$$

The simplest example is that in which each set contains only two elements and all possible combinations of the t elements, taken in pairs, appear in the different sets. This series of pairs can be written out by the experimenter, and the method of analysis is given by Yates [20].

Let us take another example; given six elements to be taken three at a time,

$$b = {}_6C_3 = \frac{6!}{3!\,3!} = 20.$$

The 20 combinations are,

123	134	*146*	*236*	*345*
124	*135*	*156*	*245*	*346*
125	136	234	246	356
126	145	235	*256*	456.

Such unreduced designs are not necessarily economical or feasible in experimental work. It is often desirable to find some less extensive configuration. In this example half of the combinations, either those in italics or the other half, fulfill the restriction that every element occur with every other element in the same number of sets. Each pair of elements occurs twice in either group of sets. Thus, a balanced incomplete block design can be based on either half of the 20 sets as well as on all 20.

2. Combinatorial methods. Combinatorial considerations of a simple nature enable us to set up necessary conditions which balanced designs must satisfy.

We have t elements arranged in b sets of k elements each; each element occurs in r sets, and each pair of elements occurs together in a set exactly λ times. Then we must have

$$tr = bk, \qquad r(k - 1) = \lambda(t - 1).$$

The first of these equations expresses the fact that the total number of plots must be equal both to the product of elements by replications and to the product of sets by number of elements per set; the second, that the number of pairs into which a given element enters must equal λ times the remaining number of elements.

It is convenient to write

$$r = \frac{\lambda(t - 1)}{k - 1}, \qquad b = \frac{\lambda t(t - 1)}{k(k - 1)}.$$

Since the numbers t, b, r, k, λ must be integers, it is easy to obtain lower limits for any three in terms of the other two.

To give a general classification, the configurations have been divided into classes according to the value of λ. Because of the practical limitations in experimentation, table I has been expanded only to include $\lambda = 6$ and the k values from 1–14. It may be well to call attention to the fact that duplications occur in the different classes of table I. For instance in the class, $\lambda = 1$, for $k = 6$, $t = 15m + 1$, and $m = 1$, then $b = 8$, and $r = 3$. In order to construct a design, the following condition is necessary; $r \geq k$ and therefore $b \geq t$. In this example, the condition is met if b, r and λ are multiplied by 2, the resulting design is $t = 16$, $b = 16$, $r = 6$, $k = 6$ and $\lambda = 2$. This configuration is a duplicate of the design in the class, $\lambda = 2$, for $k = 6$ and $m = 1$. In many of the configurations where λ is 3, 4, 5, or 6, a common factor can be cancelled from b, r and λ giving a design listed in the classes , $\lambda = 1$, 2 or 3.

It should be emphasized that the conditions under which table I was derived are necessary, but not sufficient, for the existence of a complete configuration. For example, consider the following configurations which satisfy the necessary conditions for a design.

Sub class (table I)	m	t	b	r	k	λ
$10m + 5$	1	15	21	7	5	2
$21m + 1$	1	22	22	7	7	2
$15m + 6$	2	36	42	7	6	1
$42m + 1$	1	43	43	7	7	1
$45m + 10$	2	100	110	11	10	1
$110m + 1$	1	111	111	11	11	1.

No configurations of the above specification can actually be constructed.

A selected group of configurations from table I is given in table II. Only those configurations whose k, r and λ lie within practical limits, and whose

TABLE I

Classes of Configurations

k	Class $\lambda = 1$ t	t	$\lambda = 2$ i	t	$\lambda = 3$ t	t	$\lambda = 4$ t	t	$\lambda = 5$ t	t	$\lambda = 6$ t	t
3	$6m + 1$	$6m + 3$	$3m + 1$	$3m + 3$	$2m + 1$	$2m + 3$						
4	$12m + 1$	$12m + 4$	$6m + 1$	$6m + 4$	$4m + 1$	$4m + 4$	$3m + 1$	$3m + 4$				
5	$20m + 1$	$20m + 5$	$10m + 1$	$10m + 5$	$20m + 1$	$20m + 5$	$5m + 1$	$5m + 5$	$4m + 1$	$4m + 5$		
6	$15m + 1$	$15m + 6$	$15m + 1$	$15m + 6$	$10m + 1$	$10m + 6$	$15m + 1$	$15m + 6$	$3m + 1$	$3m + 6$	$5m + 1$	$5m + 6$
7	$42m + 1$	$42m + 7$	$21m + 1$	$21m + 7$	$14m + 1$	$14m + 7$	$21m + 1$	$21m + 7$	$42m + 1$	$42m + 7$	$7m + 1$	$7m + 7$
8	$56m + 1$	$56m + 8$	$28m + 1$	$28m + 8$	$56m + 1$	$56m + 8$	$14m + 1$	$14m + 8$	$56m + 1$	$56m + 8$	$28m + 1$	$28m + 8$
9	$72m + 1$	$72m + 9$	$36m + 1$	$36m + 9$	$24m + 1$	$24m + 9$	$18m + 1$	$18m + 9$	$72m + 1$	$72m + 9$	$12m + 1$	$12m + 9$
10	$45m + 1$	$45m + 10$	$45m + 1$	$45m + 10$	$30m + 1$	$30m + 10$	$45m + 1$	$45m + 10$	$18m + 1$	$18m + 10$	$15m + 1$	$15m + 10$
11	$110m + 1$	$110m + 11$	$55m + 1$	$55m + 11$	$110m + 1$	$110m + 11$	$55m + 1$	$55m + 11$	$22m + 1$	$22m + 11$	$55m + 1$	$55m + 11$
12	$132m + 1$	$132m + 12$	$66m + 1$	$66m + 12$	$44m + 1$	$44m + 12$	$33m + 1$	$33m + 12$	$132m + 1$	$132m + 12$	$22m + 1$	$22m + 12$
13	$156m + 1$	$156m + 13$	$78m + 1$	$78m + 13$	$52m + 1$	$52m + 13$	$39m + 1$	$39m + 13$	$156m + 1$	$156m + 13$	$26m + 1$	$26m + 13$
14	$91m + 1$	$91m + 14$	$91m + 1$	$91m + 14$	$182m + 1$	$182m + 14$	$91m + 1$	$91m + 14$	$91m + 1$	$91m + 14$	$91m + 1$	$91m + 14$
.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.

$m = 1, 2, 3, \ldots .$

existence has not been disproved, have been included. The practical limits of k, r and λ, of course, are dependent upon the conditions surrounding the experiment. We have chosen to keep k within the range 3 to 10 except for a few special configurations in which t is greater than 100, in which cases k was allowed to equal 11–14. Also r has been kept within a similar limited range. (Those configurations in table II, with an asterisk preceding t, have not been constructed.)

The above limitations upon k and r give a small, selected group of configurations. However, many others either have been constructed or are known to exist. For balanced incomplete block designs, Yates [20] gives the lower limits of r for t from 4 to 25 and k from 2 to 12 but not greater than $\frac{1}{2}t$. Fisher and Yates [8] have tabulated the configurations which are known to exist having ten or less replications including all arithmetically possible configurations the existence of which has not been disproved.

Even if the existence of a configuration has not been disproved, there still remains the difficult problem of writing out the elements which are to appear in each set. Some discussion of the structure of such configurations is presented by Fisher and Yates [8] by Yates [20, 21] by Goulden [9, 10] and by Bose [4]. Additional descriptions are to follow.

While a search of the literature revealed a number of constructed configurations, yet the general theory of their formation has received relatively little consideration. The question of combinations related to the theory of configurations which is of interest here was first set forth by Kirkman [11] in 1847. He states the problem thus: "If Q_x denote the greatest number of triads that can be formed with x symbols, so that no duad shall be twice employed, then

$$3Q_x = x(x - 1)/2 - V_x$$

if for V_x we put 0, when $x = 6m + 1$ or $6m + 3$." This gives the formula for b which was given earlier in this article. Put $x = t$ and $V_x = 0$

$$b = Q_x = \frac{t(t-1)}{3 \cdot 2} = \frac{t(t-1)}{k(k-1)}.$$

Besides the theory connected with these combinatorial problems, considerable information related to the construction of the configurations has been found in the literature on finite projective geometry, especially the geometry which applies to the theory of groups.

An extensive discussion of the $\lambda = 1$ class of configurations (as listed in table I) can be found in the literature. The theory of the formation of the configurations for the sub-class $t = 6m + 3$ has been summarized by Ball [1]. This is the Kirkman "school-girl problem" for which Eckenstein [7] lists 48 papers and 5 books written during the years 1847–1911 dealing with this subject. The problem was first published in the Lady's and Gentleman's Diary for 1850 [12]. It is usually stated that "a schoolmistress was in the habit of taking her girls for a daily walk. The girls were fifteen in number, and were arranged in five rows of three each, so that each girl might have two companions. The problem

is to dispose of them so that for seven consecutive days no girl will walk with any of her school-fellows in any triplet more than once." For this particular subclass ($t = 6m + 3$, $k = 3$), this type of configuration has been shown to exist

TABLE II

Selected Group of Configurations

(Balanced Incomplete Block Designs)

t	b	r	k	λ
7	7	3	3	1 Y.S.[1]
7	7	4	4	2
8	14	7	4	3
9	12	4	3	1
9	6 + 6	2	3	1 L.S.[2]
9	18	8	4	3
9	18	10	5	5
9	12	8	6	5
10	30	9	3	2
10	15	6	4	2
10	18	9	5	4
10	15	9	6	5
11	11	5	5	2
11	11	6	6	3
13	26	6	3	1
13	13	4	4	1 Y.S.
13	13	9	9	6
15	35	7	3	1
15	15	7	7	3
15	15	8	8	4
16	20	5	4	1
16	20 + 20	5	4	2 L.S.
16	16	6	6	2
16	16	10	10	6
19	57	9	3	1
19	19	9	9	4
19	19	10	10	5
21	70	10	3	1
21	21	5	5	1 Y.S.
*21	28	8	6	2
*21	30	10	7	3
*25	50	8	4	1
25	30	6	5	1
25	15 + 15	3	5	1 L.S.
*25	25	9	9	3
28	63	9	4	1
28	36	9	7	2
*29	29	8	8	2
31	31	6	6	1 Y.S.
*31	31	10	10	3
*36	45	10	8	2
37	37	9	9	2
*41	82	10	5	1
*46	69	9	6	1
*46	46	10	10	2
49	56	8	7	1
49	28 + 28	4	7	1 L.S.
*51	85	10	6	1
57	57	8	8	1 Y.S.
64	72	9	8	1
64	72 + 72	9	8	2 L.S.
73	73	9	9	1 Y.S.
81	90	10	9	1
81	45 + 45	5	9	1 L.S.
91	91	10	10	1 Y.S.
121	132	12	11	1
121	66 + 66	6	11	1 L.S.
133	133	12	12	1 Y.S.
169	182	14	13	1
169	91 + 91	7	13	1 L.S.
183	183	14	14	1 Y.S.

* Have not been constructed.
[1] Youden squares.
[2] Lattice squares.

for every possible value of t. Most of the solutions were worked by H. E. Dudeney and O. Eckenstein. They are given by Ball [1] for all t's less than 100, that is, for t = 9, 15, 21, 27, 33, 39, 45, 51, 57, 63, 69, 75, 81, 87, 93 and 99. Ball describes several methods of constructing such configurations, as cycles, combinations of cycles, scalene triangles inscribed in the circle, focal and analyti-

cal methods. As an illlustration of the school-girl problem, the construction of the configuration for $t = 9$, $b = 12$, $r = 4$, $k = 3$ and $\lambda = 1$ will be shown. Scalene triangles are inscribed in a circle with certain specifications (to be fulfilled) giving the three sets of triplets for the first day as follows,

Set	Group I		
(1)	k	1	5
(2)	3	4	6
(3)	7	8	2.

By rotation or by cyclic substitution the other three groups are secured:

Set	Group II				Group III				Group IV		
(4)	k	2	6	(7)	k	3	7	(10)	k	4	8
(5)	4	5	7	(8)	5	6	8	(11)	6	7	1
(6)	8	1	3,	(9)	1	2	4,	(12)	2	3	5.

Then placing $k = 9$, we have the configuration for $t = 9$, $b = 12$, and $r = 4$. Note that in the school-girl problem the sets are grouped into complete replications of the elements. This problem of 9 girls taken 3 at a time has been subjected to an exhaustive examination. There are 840 arrangements but only one fundamental solution. In the case of 15 girls, the number of fundamental solutions according to Mulden [14] and Cole [6], is seven. Ball mentions the Kirkman problem in quartets which is the sub-class $t = 12m + 4$, for $k = 4$. He states that this has been solved for cases where m does not exceed 49. He also states, "I conjecture that similar methods are applicable to corresponding problems about quintets, sextets, etc."

Before leaving the school-girl problem, an illustration will be given of $t = 28$, $b = 63$, $r = 9$, $k = 4$ and $\lambda = 1$. The following framework was set up by Dr. C. P. Winsor using suggestions from Netto [15].

k	a	b	c
a_1	a_8	b_3	b_6
a_2	a_7	b_1	b_8
a_3	a_6	c_4	c_5
a_4	a_5	c_1	c_8
b_2	b_7	c_3	c_6
b_4	b_5	c_2	c_7.

a, b and c each have every internal difference once and only once; and each pair a-b, a-c and b-c must have every external difference once and only once. The nine groups are given in table III. The cyclic substitution is within three sets, a, b and c. That is,

in group I, $a = 1,\quad a_1 = 2,\quad a_2 = 3,\quad \cdots,\quad a_8 = 9;$
in group II, $a = 2,\quad a_1 = 3,\quad a_2 = 4,\quad \cdots,\quad a_8 = 1;$
in group III, $a = 3,\quad a_1 = 4,\quad a_2 = 5,\quad \cdots,\quad a_8 = 2;$
etc.

Netto [15] discusses t elements in sets of k, every set of 2 elements to occur together in a set exactly λ times. He deals with $\lambda = 1$, and gives a discussion of both sub-classes when $k = 3$, that is, for $t = 6m + 1$ and $t = 6m + 3$. Reiss [16] and Moore [13] have proved that configurations can be constructed for all values of t if $k = 3$. This is the type of information which is valuable in answer-

TABLE III

Configuration for $t = 28$, $b = 63$, $r = 9$, $k = 4$, $\lambda = 1$

				Group I				Group II				Group III				Group IV			
k	a	b	c	28	1	10	19	28	2	11	20	28	3	12	21	28	4	13	22
a_1	a_8	b_3	b_6	2	9	13	16	3	1	14	17	4	2	15	18	5	3	16	10
a_2	a_7	b_1	b_8	3	8	11	18	4	9	12	10	5	1	13	11	6	2	14	12
a_3	a_6	c_4	c_5	4	7	23	24	5	8	24	25	6	9	25	26	7	1	26	27
a_4	a_5	c_1	c_8	5	6	20	27	6	7	21	19	7	8	22	20	8	9	23	21
b_2	b_7	c_3	c_6	12	17	22	25	13	18	23	26	14	10	24	27	15	11	25	19
b_4	b_5	c_2	c_7	14	15	21	26	15	16	22	27	16	17	23	19	17	18	24	20

Group V				Group VI				Group VII				Group VIII				Group IX			
28	5	14	23	28	6	15	24	28	7	16	25	28	8	17	26	28	9	18	27
6	4	17	11	7	5	18	12	8	6	10	13	9	7	11	14	1	8	12	15
7	3	15	13	8	4	16	14	9	5	17	15	1	6	18	16	2	7	10	17
8	2	27	19	9	3	19	20	1	4	20	21	2	5	21	22	3	6	22	23
9	1	24	22	1	2	25	23	2	3	26	24	3	4	27	25	4	5	19	26
16	12	26	20	17	13	27	21	18	14	19	22	10	15	20	23	11	16	21	24
18	10	25	21	10	11	26	22	11	12	27	23	12	13	19	24	13	14	20	25

ing the first question in the introduction of this article; "what configurations exist?" Carmichael [5] mentions the quadruple systems $6m + 2$ and $6m + 4$ and states that the general problem of their existence appears not to have been solved. Also for the higher values of k there seems to be very little known of any generality, but it is known that for $k > 3$ there are certain configurations which are not possible.

3. **The method of geometrical configuration.** Another aid in the construction of balanced incomplete block designs is found in some of the finite projective geometries. These are described by Carmichael [5]. A tactical configuration of rank two is defined as a combination of l elements into m sets, each set containing λ distinct elements, and each element occurring in μ distinct sets,

$l = (t) =$ number of points in the geometry,
$m = (b) =$ number of lines,
$\lambda = (k) =$ number of points,
$\mu = (r) =$ number of lines on a point.

The series of finite projective geometries $PG(\kappa, p^n)$ for $\kappa > 1$ furnishes a certain infinite class of these tactical configurations. The following list gives those which have been incorporated in the list (table II) of useful balanced incomplete block designs.

Two dimensional space, $PG(2, p^n)$

p^n	$l(t)$	$m(b)$	$\lambda(k)$	$\mu(r)$
2	7	7	3	3
3	13	13	4	4
2^2	21	21	5	5
5	31	31	6	6
7	57	57	8	8
2^3	73	73	9	9
3^2	91	91	10	10
11	133	133	12	12
13	183	183	14	14.

Three dimensional space, $PG(3, p^n)$

p^n	l	m	λ	μ
2	15	35	7	3.

From the Euclidean geometry $EG(\kappa, p^n)$ for $\kappa > 1$ other tactical configurations can be constructed. These are formed from the $PG(\kappa, p^n)$ by omitting a given line from the two dimensional space and a plane from the three dimensional space configurations. Some of the resulting designs are:

Two dimensional space, $EG(2, p^n)$

p^n	l	m	λ	μ
2	4	6	3	2
3	9	12	4	3
2^2	16	20	5	4
5	25	30	6	5
7	49	56	8	7
2^3	64	72	9	8
3^2	81	90	10	9
11	121	132	12	11
13	169	182	14	13.

Methods are available for constructing the two dimensional space $PG(\kappa, p^n)$ and the corresponding $EG(\kappa, p^n)$ configurations where p is a prime number. This being true, we can also construct the completely orthogonalized squares from the $EG(\kappa, p^n)$ geometry. The reverse situation in which these configurations are constructed by using the completely orthogonalized squares is to be illustrated. These squares consist of superimposed Latin squares, fulfilling the condition that each number from the second Latin square occurs once and only once with each number in the first Latin square. As an example take the two Latin squares:

Latin Square I				Latin Square II		
1	2	3		1	3	2
2	3	1		2	1	3
3	1	2,		3	2	1.

Superimpose square II upon square I to get the completely orthogonalized 3 x 3 square,

11	23	32
22	31	13
33	12	21.

The first number in each cell is a value from square I; the second number in each cell is from square II. Note that the numbers in the second place in each cell occur once and only once with each of the first numbers, that is 1-1, 1-3, and 1-2. The completely orthogonalized squares have been proven to exist for all prime numbers and for powers of prime numbers. The solution of this problem was secured independently by Bose [2] and by Stevens [18]. Those of sides 2, 2^2, 2^3, 2^4, 2^5, 2^6, 3, 3^2, 3^3, 3^4, 5, 5^2, 5^3, 7, 7^2, 11 and 13 have been given.

The completely orthogonalized 3 x 3 square may be used to construct

11	*1*	23	*4*	32	*7*
22	*2*	31	*5*	13	*8*
33	*3*	12	*6*	21	*9*

a balanced incomplete block design. The italic numbers, which follow the cell numbers, designate the 9 elements which are to be arranged in four groups of three sets. Group I is formed by placing the elements from each row into separate sets, in group II the elements from the three columns are placed in three sets; in group III the first set (7) consists of the elements which follow 1 in the first place in the cells, set (8) consists of the elements which follow 2 in the first place in the cells; and group IV is assembled in the same way as group III except the numbers in the second place in the cells are used to select the elements for each set. Thus we have the configuration:

Set	Group I (rows)				Group II (columns)				Group III (first place)				Group IV (second place)		
(1)	1	4	7	(4)	1	2	3	(7)	1	6	8	(10)	1	5	9
(2)	2	5	8	(5)	4	5	6	(8)	2	4	9	(11)	2	6	7
(3)	3	6	9	(6)	7	8	9	(9)	3	5	7	(12)	3	4	8

In the 12 sets of 3 elements, each of the 9 elements occurs with every other element once and only once in a set.

This is an illustration of one series of configurations which can be constructed with the aid of the completely orthogonalized squares. These are the $EG(\kappa, p^n)$ in two dimensional space when $\kappa = 2$ and $p^n = 2, 3, 2^2, 5, 7, 2^3, 3^2, 11, 13, \ldots$ The $PG(\kappa, p^n)$ configurations can be written by adding $(k + 1)$ elements to the previous group of configurations. For example, the elements 10, 11, 12 and 13 may be added to the groups, one to each group. That is, 10 is added to each set in group I, 11 is added to each set in group II, 12 to group III and 13 to group IV. An additional set must be added to include these four new elements. A configuration for $t = 13, b = 13, k = 4, r = 4$ and $\lambda = 1$ results.

Set																			
(1)	1	4	7	10	(4)	1	2	3	11	(7)	1	6	8	12	(10)	1	5	9	13
(2)	2	5	8	10	(5)	4	5	6	11	(8)	2	4	9	12	(11)	2	6	7	13
(3)	3	6	9	10	(6)	7	8	9	11	(9)	3	5	7	12	(12)	3	4	8	13
															(13)	10	11	12	13.

The 13 sets are made up of 4 elements each. These designs are symmetrical for sets and elements, that is, every pair of elements occurs together in the same number of sets, also, every pair of sets has the same number of elements in common. Discussion of the construction of these designs with illustrations are given in references [20, 8, 9] and [19].

In the $PG(\kappa, p^n)$ series of designs, as constructed by means of completely orthogonalized squares, the sets cannot be arranged in replication groups. However, these configurations can be arranged in Youden squares [22] in which all the sets are placed side by side and all the elements in a single row form a complete replication. This method of arrangement has been of considerable value in experimentation with plants. The Youden squares are the $PG(\kappa, p^n)$ when $\kappa = 2$. Singer [17] gives a partial list of the (reduced) perfect difference sets (table IV), only a single set for each p^n. The number of distinct perfect difference sets (or the number of distinct perfect partitions) for a given p^n is equal to $\varphi(q)/3n$. Since each perfect difference set can be paired with its inverse, the number is even.

The construction of one of the Youden squares from its perfect difference set will be illustrated. Consider $p^n = 3$ then $q = p^{2n} + p^n + 1 = 3^2 + 3 + 1 = 13$. There are two perfect difference sets with their inverses for $q = 13$. One perfect difference set is 0, 1, 3, 9 which has the perfect partition 1, 2, 6, 4 which will add in succession to each number from 1 to and including 13, and also 1, 2, 6, 4

add to 13. The elements of the perfect difference set are put in set (1) except that 13 replaces 0. Set (2) is secured by a one-step cyclic substitution, 1 for 13, 2 for 1, 4 for 3 and 10 for 9. This process is continued until there are thirteen sets. If the substitution is applied to set (13), the elements in set (1) are secured.

		Set												
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Replica-	A	13	1	2	3	4	5	6	7	8	9	10	11	12
tion	B	1	2	3	4	5	6	7	8	9	10	11	12	13
	C	3	4	5	6	7	8	9	10	11	12	13	1	2
	D	9	10	11	12	13	1	2	3	4	5	6	7	8.

This is the Youden square for $t = 13$, $b = 13$, $r = 4$, $k = 4$, and $\lambda = 1$. The elements in each row form a complete replication.

TABLE IV

Singer's list of perfect difference sets

p^n	q	$\frac{\varphi(q)}{3^n}$	Perfect difference set																
2	7	2	0	1	3														
2^2	21	2	0	1	4	14	16												
2^3	73	8	0	1	3	7	15	31	36	54	63								
2^4	273	12	0	1	3	7	15	31	63	90	116	127	136	181	194	204	233	238	255
3	13	4	0	1	3	9													
3^2	91	12	0	1	3	9	27	49	56	61	77	81							
5	31	10	0	1	3	8	12	18											
7	57	12	0	1	3	13	32	36	43	52									
11	133	36	0	1	3	12	20	34	38	81	88	94	104	109					
13	183	40	0	1	3	16	23	28	42	76	82	86	119	137	154	175			

$$t = q = p^{2n} + p^n + 1$$

A third series of configurations, called Lattice squares or quasi-Latin squares [21] can be constructed by using the completely orthogonalized squares. The groups of sets on page 78 are taken in pairs. For each pair a square is constructed having its rows formed by the sets of one group and its columns by the sets of another group. For example, square I below is made so that the sets of group I form the rows and the sets of group II form the columns. Square II is the combination of groups III and IV.

Square I

1	4	7
2	5	8
3	6	9

Square II

1	6	8
9	2	4
5	7	3

In this lattice square each pair of elements occurs together once only in either a row or a column of either one of the squares. Also, every element occurs with every other element once in one column and one row from each square.

A device known as "complements" gives several configurations. From an arrangement having $k \neq \frac{1}{2}t$, a second one can be obtained for the same number of elements, in sets of $t - k$ units. This is done by replacing each set by its complement, that is, by a set containing all the elements missing from the original set. An illustration follows:

$t = 7, \quad b = 7$
$r = 3, \quad k = 3$
$\lambda = 1$

Set			
(1)	1	2	4
(2)	2	3	5
(3)	3	4	6
(4)	4	5	7
(5)	5	6	1
(6)	6	7	2
(7)	7	1	3,

$t = 7, \quad b = 7$
$r = 4, \quad k = 4$
$\lambda = 2$

Set				
(1)	3	5	6	7
(2)	1	4	6	7
(3)	1	2	5	7
(4)	1	2	3	6
(5)	2	3	4	7
(6)	1	3	4	5
(7)	2	4	5	6.

While the triple systems, quadruple systems, etc., which have been considered by some mathematicians, do furnish designs meeting the balance requirements, they are usually not suitable for experimental purposes. A quadruple system requires that every possible triple of elements occur once and only once together in a block. Since we need only every pair together once ($\lambda = 1$) or more, only the triple systems are generally useful.

4. **Summary.** The mathematical theory of configuration has been helpful in the construction of the balanced incomplete block designs. It would be useful to know (a) what configurations (within the useful range) exist, (b) how these configurations may be constructed. In table I the configurations have been classified according to the value of λ, while in table II configurations within a useful range have been listed. Of the designs in this table which have not been constructed, some are known to exist. Those aids which have been used in the construction of the balanced incomplete block designs have been briefly discussed.

REFERENCES

[1] Ball, W. W. R. Revised by Coxeter, H. S. M. *Mathematical recreations and essays*, Macmillan and Co., London, 11th Edition, 1939.

[2] Bose, R. C. "On the Application of the Properties of Galois Fields to the Problem of the Construction of Hyper-Graeco Latin Squares," *Sankhyā*, Vol. 3(1938), pp. 323–338.

[3] Bose, R. C. and Nair, K. R. "Partially Balanced Incomplete Block Designs," *Sankhyā*, Vol. 4(1939), pp. 337–372.

[4] BOSE, R. C. "On the Construction of Balanced Incomplete Block Designs," *Ann. of Eugenics*, Vol. 9 Part 4 (in the press).

[5] CARMICHAEL, R. D. *Introduction to the theory of groups of finite order*, Ginn and Company, 1937.

[6] COLE, F. N. "Kirkman Parades," *Bull. of the Am. Math. Society*, Vol. 28(1922), pp. 435-437.

[7] ECKENSTEIN, O. "Bibliography of Kirkman's School-girl Problem," *Messenger of Mathematics*, Vol. 41(1911), pp. 33-36.

[8] FISHER, R. A. AND YATES, F. *Statistical tables for biological, agricultural, and medical research*, Oliver and Boyd, Edinburgh, 1938.

[9] GOULDEN, C. H. "Modern Methods for Testing a Large Number of Varieties," *Dominion of Canada, Department of Agriculture, Tech. Bull.* 9, 1937.

[10] GOULDEN, C. H. *Methods of statistical analysis*, John Wiley and Sons, New York, 1939.

[11] KIRKMAN, T. P. "On a Problem in Combinations," *Cambridge and Dublin Math. Jour.*, Vol. 2(1847), pp. 191-204.

[12] KIRKMAN, T. P. "Query," *Lady's and Gentleman's diary*, p. 48, 1850.

[13] MOORE, E. H. "Concerning Triple Systems," *Math. Ann.*, Vol. 43(1893), pp. 271-285.

[14] MULDER, P. *Kirkman-Systemen* (Groningen Dissertation), Leiden, 1917.

[15] NETTO, EUGEN. *Lehrbuch der combinatorik*. Verlag und Druck von B. G. Teubner, Leipzig und Berlin, 1927. Zweite Auflage.

[16] REISS, M. "Ueber eine Steinersche combinatorische Aufgabe, welche im 45sten Bande dieses Journals, Seite 181, gestellt worden ist," *Crelles Journal*, Vol. 56(1859), pp. 326-344.

[17] SINGER, JAMES. "A Theorem in Finite Projective Geometry and some Applications to Number Theory," *Trans. Am. Math. Society*, Vol. 43(1938), pp. 377-385.

[18] STEVENS, W. L. "The Completely Orthogonalized Latin Square," *Ann. of Eugenics*, Vol. 9(1939), pp. 82-93.

[19] WEISS, MARTIN G. AND COX, GERTRUDE M. "Balanced Incomplete Block and Lattice Square Designs for Testing Yield Differences Among Large Numbers of Soybean Varieties," *Iowa Agr. Exp. Sta. Res. Bull.*, 257(1939), pp. 293-316.

[20] YATES, F. "Incomplete Randomized Blocks," *Ann. of Eugenics*, Vol. 7(1936), pp. 121-140.

[21] YATES, F. "A Further Note on the Arrangement of Variety Trials: quasi-Latin Squares," *Ann. of Eugenics*, Vol. 7(1937), pp. 319-332.

[22] YOUDEN, W. J. "Use of Incomplete Block Replications in Estimating Tobacco-mosaic Virus," *Contributions from Boyce Thompson Inst.*, Vol. 9(1937), pp. 41-48.

IOWA STATE COLLEGE,
AMES, IOWA.

A COMPARISON OF ALTERNATIVE TESTS OF SIGNIFICANCE FOR THE PROBLEM OF m RANKINGS[1]

By Milton Friedman

A paper published in 1937 [2] suggested that the consilience of a number of sets of ranks can be tested by computing a statistic designated χ_r^2. A mathematical proof by S. S. Wilks demonstrated that the distribution of χ_r^2 approaches the ordinary χ^2 distribution as the number of sets of ranks increases. The rapidity with which this limiting distribution is approached was investigated by obtaining the exact distributions of χ_r^2 for a number of special cases. It was concluded that "when the number of sets of ranks is moderately large (say greater than 5 for four or more ranks) the significance of χ_r^2 can be tested by reference to the available χ^2 tables" [2, p. 695]. The use of the normal distribution was recommended when the number of ranks in each set is large, but the number of sets of ranks is small, although no rigorous justification of this procedure was presented.

Except for the few special cases for which exact distributions were given, the paper did not provide a test of significance for data involving less than six sets of ranks and a small or moderate number of ranks in each set. This important gap has now been filled by M. G. Kendall and B. Babington Smith [1]. In addition, they furnish a somewhat more exact test of significance for tables of ranks for which the earlier article recommended the use of the χ^2 distribution.

Kendall and Smith use a different statistic, W, defined as χ_r^2 divided by its maximum value, $m(n - 1)$, where n is the number of items ranked, and m the number of sets of ranks.[2] The new statistic (independently suggested by W. Allen Wallis [3] who terms it the rank correlation ratio and denotes it by η_r^2) is thus not fundamentally different from χ_r^2. A more radical innovation is the improvement in the test of significance that they suggest. Instead of testing χ_r^2 by reference to the χ^2 distribution for $n - 1$ degrees of freedom, Kendall and Smith, generalizing from the first four moments of W, recommend that the significance of W be tested by reference to the analysis of variance distribution (Fisher's z-distribution) with $z \doteq \frac{1}{2}\log_e\left(\frac{(m-1)W}{1-W}\right)$, $n_1 = (n-1) - \frac{2}{m}$, $n_2 = (m-1)\left[(n-1) - \frac{2}{m}\right]$. For small values of m and n, they introduce con-

[1] The author is indebted to Mr. W. Allen Wallis for valuable criticism and to Miss Edna R. Ehrenberg for computational assistance.

[2] This is Kendall and Smith's notation which will be used in the present paper. The original paper [2] designated the number of items ranked by p, and the number of sets of ranks by n.

tinuity corrections, substituting for $W = \dfrac{12S}{m^2(n^3 - n)}$, the statistic

$$W_c = \frac{S - 1}{\dfrac{m^2(n^3 - n)}{12} + 2} = \frac{W - \dfrac{12}{m^2(n^3 - n)}}{1 + \dfrac{24}{m^2(n^3 - n)}},$$

where S is the observed sum of squares of the deviations of sums of ranks from the mean value, $m(n + 1)/2$. Comparison with exact distributions of W (or S) for special cases indicates that this test yields very good approximations to the correct probabilities.

In the limit the two tests of significance are identical. Neglecting the correction for continuity, $z = \dfrac{1}{2}\log_e\left(\dfrac{(m-1)\chi_r^2}{m(n-1) - \chi_r^2}\right) \rightarrow \dfrac{1}{2}\log_e\left(\dfrac{\chi_r^2}{n-1}\right)$, $n_2 = (m - 1)\left[(n - 1) - \dfrac{2}{m}\right] \rightarrow \infty$, and $n_1 = (n - 1) - \dfrac{2}{m} \rightarrow (n - 1)$ as $m \rightarrow \infty$. For $n_2 = \infty$, the analysis of variance distribution is identical with the distribution of $\frac{1}{2}\log_e \dfrac{\chi^2}{n_1}$. The difference between the two tests is thus that one, χ^2, uses a single (limiting) distribution for all values of m, whereas the other, z, adapts the distribution to the value of m.

The necessity of taking into account the value of m, while it increases the flexibility of the distribution, makes the z test somewhat less convenient in practice than the χ^2 test. Additional computation is required to obtain the values of n_1 and n_2, and to make the continuity corrections. It is also fairly laborious to test the significance of the result, if exact values of z at any level of significance are required. In these instances, two-way interpolation of reciprocals in the analysis of variance tables is necessary since both n_1 and n_2 are always fractional. These difficulties make it desirable to investigate the rapidity with which the significance levels given by the z test approach those given by the χ^2 test, and thus determine the range of values of m and n for which the simpler test can safely be employed. This investigation will yield as a by product the .05 and .01 significance values of χ_r^2 (or W or S) for selected values of m and n as determined by the z test.

Table I presents a summary comparison of the values of χ_r^2 at the .05 and .01 levels of significance as shown by (1) exact distributions, (2) the z test with continuity corrections, (3) the χ^2 test.[3] The significance values are expressed in terms of χ_r^2 rather than W because, for a given number of ranks per set (i.e., a given n), the significance values given by the χ^2 test are the same regardless of the number of sets of ranks (i.e., of the value of m). This would not be so if W were employed, since $W = \chi_r^2/m\,(n - 1)$. The expected value of W depends on

[3] The values of χ_r^2 computed using the z test that are given in Tables I and II were obtained with the aid of Fisher and Yates' Table V [4]. Linear interpolation of reciprocals was employed throughout.

m and approaches zero as $m \to \infty$ while the expected value of χ_r^2 is equal to $n - 1$ for all values of m.

The values given by the z test agree remarkably well with the exact values. With but two exceptions (the .01 values for $n = 3$, $m = 8$ and 10) the exact value differs very much less from the value given by the z test than from the value given by the χ^2 test. In all but three of the 12 comparisons, the z test gives a value below the correct one.[4]

TABLE I

Comparison of Values of χ_r^2 at .05 and .01 Levels of Significance Yielded by Exact Distributions, z Test with Continuity Corrections, and χ^2 Test

n	m	.05 Level of Significance				.01 Level of Significance			
		From Exact Distribution		From z test with continuity corrections	From χ^2 test	From Exact Distribution		From z test with continuity corrections	From χ^2 test
		Limits	Interpolated value*			Limits	Interpolated value*		
3	8	5.25-6.25	6.16	6.012	5.991		9.00	8.35	9.21
	9	6.0 -6.22	6.17	6.004	5.991		8.67	8.44	9.21
	10	5.6 -6.2	6.08	5.999	5.991	8.6 - 9.6	9.04	8.51	9.21
	∞			5.991	5.991			9.21	9.21
4	4	7.5 -7.8	7.54	7.43	7.82	9.3 - 9.6	9.42	9.21	11.34
	5	7.32-7.8	7.54	7.52	7.82	9.72- 9.96	9.87	9.66	11.34
	6	7.4 -7.6	7.49	7.57	7.82		10.00	9.95	11.34
	∞			7.82	7.82			11.34	11.34
5	3	8.27-8.53	8.41	8.59	9.49	9.87-10.13	10.05	10.08	13.28
	∞			9.49	9.49			13.28	13.28

* Computed by linear interpolation of probabilities.

Table II gives for a very much larger number of values of m and n the .05 and .01 values of χ_r^2 computed on the basis of the z test with continuity correc-

[4] These comparisons duplicate some of those made by Kendall and Smith and merely serve to confirm their conclusion that the z test with continuity corrections gives exceedingly good results.

The values obtained using the z test without continuity corrections agree less well with the exact values than those obtained with the aid of the continuity corrections. However even if no continuity corrections are made the z test in general yields values closer to the exact values than does the χ^2 test.

TABLE II

Values of χ_r^2 at .05 and .01 Levels of Significance Computed on the Basis of Kendall and Smith's z test, with Continuity Corrections; .10, .075, .02, .015 Values of χ^2

m	n: 3	4	5	6	7
	Values at .05 Level of Significance				
3			8.59	9.90	11.24
4		7.43	8.84	10.24	11.62
5		7.52	8.98	10.42	11.84
6		7.57	9.08	10.54	11.97
8	6.012	7.63	9.18	10.68	12.14
10	5.999	7.67	9.25	10.76	12.23
15	5.985	7.72	9.33	10.87	12.36
20	5.983	7.74	9.37	10.92	12.42
100	5.987	7.80	9.46	11.04	12.56
∞	5.991	7.82	9.49	11.07	12.59
χ^2 (.10)	4.605	6.25	7.78	9.24	10.64
χ^2 (.075)*	5.18	6.90	8.49	10.00	11.45
	Values at .01 Level of Significance				
3			10.08	11.69	13.26
4		9.21	10.93	12.59	14.19
5		9.66	11.42	13.11	14.74
6		9.95	11.74	13.45	15.09
8	8.35	10.31	12.13	13.87	15.53
10	8.51	10.52	12.37	14.11	15.79
15	8.74	10.79	12.67	14.44	16.14
20	8.85	10.93	12.82	14.60	16.31
100	9.14	11.26	13.19	14.99	16.71
∞	9.21	11.34	13.28	15.09	16.81
χ^2 (.02)	7.82	9.84	11.67	13.39	15.03
χ^2 (.015)*	8.40	10.46	12.34	14.09	15.77

* Computed from Fisher and Yates' Table IV (4) by linear interpolation between the logarithms of the probabilities.

tions. The values entered for $m = \infty$ are obtained from χ^2 tables for $n - 1$ degrees of freedom and are the significance values by the χ^2 test for all values of m. It is apparent that as m increases the .01 and .05 values of χ_r^2 approach their limiting values very rapidly. For $n = 7$, two-thirds of the difference between the .05 values for $m = 3$ and $m = \infty$, and an even larger proportion of the difference between the .01 values, disappears by the time $m = 10$; and the situation is similar for the other values of n. Except for the .05 values for $n = 3$, the approach to the limit is monotonic from below. The use of the χ^2 test thus tends to lead to the overestimation of the significance values and of the probabilities attached to observed values of χ_r^2. It is clear, however, that for large and even moderate values of m the χ^2 test is, for all practical purposes, equivalent to the z test.

In order to determine more precisely the range of values of m and n for which the approximation given by the χ^2 test is adequate, it is necessary to adopt some convention about the error in estimated significance values of χ_r^2 that is tolerable. Since the conclusion drawn from an observed χ_r^2 depends on the probability that it will be exceeded by chance, this convention clearly should be expressed in terms of the error in the probability.

The structure of published χ^2 tables makes it convenient to accept an estimated probability between .10 and .05 as a tolerable approximation to a correct probability of .05, and an estimated probability between .02 and .01 as a tolerable approximation to a correct probability of .01. These ranges of tolerance are entirely on one side of the correct probability because, as pointed out above, the error in using the χ^2 test is consistent in direction. These ranges are purely arbitrary, of course, and many may think them too broad.

On the basis of this or some similar convention it is possible to make objective statements concerning the range of values of m and n for which the χ^2 test is adequate. The next to the last line in the first section of Table II gives the .10 values of χ^2; the next to the last line in the second section, the .02 values. All the .05 values of χ_r^2 shown in the table exceed the .10 value of χ^2. Using the χ^2 test, all of the values (with two exceptions for $n = 3$) would signify a probability greater than .05 but less than .10. Thus the error made at the .05 level is within the admissible range according to the suggested convention. The χ^2 test is therefore an adequate substitute for the z test at the .05 level for all values of m and n except possibly for a few of the values for which exact distributions are available.

As might be expected, the χ^2 test is less satisfactory at the .01 level. For values of m less than six, the .01 values of χ_r^2 computed using the z test with continuity corrections are less than the .02 value of χ^2. For m greater than 5, the values of χ_r^2 in the table would all be accorded a probability greater than .01 but less than .02 if the χ^2 test were employed. As already noted, this is the range of values of m for which the original paper suggested the χ^2 test could validly be used [2, p. 695].

In view of the arbitrary nature of the convention as to the permissible error

in the probability attached to an observed value of χ_r^2, it is interesting to investigate the effect of an alternative and stricter convention, namely, that only probabilities from .075 to .05 and from .015 to .01 be accepted as approximations to correct probabilities of .05 and .01 respectively. The .075 and .015 values of χ^2 are given in the last lines of the two sections of Table II. On the basis of this convention the χ^2 test is adequate at the .05 level for m greater than three, and

TABLE III

Values of S at .05 and .01 Levels of Significance Computed on the Basis of Kendall and Smith's z test, with Continuity Corrections

m	n					Additional values for $n = 3$	
	3	4	5	6	7	m	S
			Values at .05 Level of Significance				
3			64.4	103.9	157.3	9	54.0
4		49.5	88.4	143.3	217.0	12	71.9
5		62.6	112.3	182.4	276.2	14	83.8
6		75.7	136.1	221.4	335.2	16	95.8
8	48.1	101.7	183.7	299.0	453.1	18	107.7
10	60.0	127.8	231.2	376.7	571.0		
15	89.8	192.9	349.8	570.5	864.9		
20	119.7	258.0	468.5	764.4	1158.7		
			Values at .01 Level of Significance				
3			75.6	122.8	185.6	9	75.9
4		61.4	109.3	176.2	265.0	12	103.5
5		80.5	142.8	229.4	343.8	14	121.9
6		99.5	176.1	282.4	422.6	16	140.2
8	66.8	137.4	242.7	388.3	579.9	18	158.6
10	85.1	175.3	309.1	494.0	737.0		
15	131.0	269.8	475.2	758.2	1129.5		
20	177.0	364.2	641.2	1022.2	1521.9		

at the .01 level for m greater than nine, except possibly for a few of the values for which exact distributions are available. Thus even so drastic a lowering of the permissible margin of error as halving it limits only slightly the range of values of m for which the χ^2 test is adequate.

Table II provides, of course, a direct means of testing the significance of observed values of χ_r^2 for the tabled values of m and n. For this purpose, however, Table III, giving the significance values of S is more useful, since it obviates

the necessity of converting S into χ_r^2. For $n = 3$ Table III includes a few values of m in addition to those in Table II.

SUMMARY

The preceding analysis suggests that the χ^2 test of the significance of χ_r^2 (or W or η_r^2), while less accurate than the z test proposed by Kendall and Smith, is adequate for practical purposes at the .01 level of significance if the number of sets of ranks (m) is greater than 5; and at the .05 level for any number of sets of ranks, provided the number of ranks in each set (n) is more than 3. Exact distributions are now available for $n = 3$, $m = 3$ to 10; $n = 4$, $m = 3$ to 6; $n = 5$, $m = 3$ [1]. The .05 and .01 values of χ_r^2 and S, computed using the Kendall and Smith z test with continuity corrections, are given in Tables II and III of the present note for $n = 3$ to 7 and selected values of m from 3 to 100. For n greater than 7 and m less than 6, the z test with continuity corrections should be employed. For all other combinations of n and m not covered by the exact distributions or by Tables II and III, the χ^2 test is adequate.

REFERENCES

[1] M. G. Kendall and B. Babington Smith, "The Problem of m Rankings," *Annals of Math. Stat.*, Vol. X (1939), pp. 275–87.

[2] Milton Friedman, "The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance," *Jour. Am. Stat. Assn.*, Vol. XXXII (1937), pp. 675–701.

[3] W. Allen Wallis, "The Correlation Ratio for Ranked Data," *Jour. Am. Stat. Assn.*, Vol. XXXIV (1939), pp. 533–8.

[4] R. A. Fisher and F. Yates, *Statistical Tables for Biological, Agricultural and Medical Research* (London: Oliver and Boyd, 1938).

NATIONAL BUREAU OF ECONOMIC RESEARCH,
NEW YORK.

NOTES

This section is devoted to brief research and expository articles, notes on methodology and other short items.

NOTE ON AN APPROXIMATE FORMULA FOR THE SIGNIFICANCE LEVELS OF Z

By W. G. Cochran

1. **Introduction.** An important part has been played in modern statistical analysis by the distribution of $z = \frac{1}{2} \log \frac{s_1^2}{s_2^2}$, when s_1^2 and s_2^2 are two independent estimates of the same variance. In particular, all tests of significance in the analysis of variance and in multiple regression problems are based on this distribution. Complete tabulation of the frequency distribution of z is a heavy task, because the distribution is a two-parameter one, the parameters being the number of degrees of freedom, n_1 and n_2 in the estimates s_1^2 and s_2^2. Thus each significance level of z requires a separate two-way table. Fisher constructed a table of the 5 percent points in 1925 [1], and this has since been extended by several workers [2] to the 20, 1, and 0.1 percent level for a somewhat wider range of values of n_1 and n_2.

With his original table, Fisher gave an approximate formula for the 5 percent values of z, for high values of n_1 and n_2 outside the limits of his table. The formula reads:

$$z \text{ (5 percent)} = \frac{1.6449}{\sqrt{h-1}} - 0.7843\left(\frac{1}{n_1} - \frac{1}{n_2}\right), \tag{1}$$

$$\text{where } \frac{2}{h} = \frac{1}{n_1} + \frac{1}{n_2}.$$

The constant 1.6449 is the 5 percent significance level for *a single tail* of the normal distribution, and the constant 0.7843 will be found to be $\frac{1}{6}\{2 + (1.6449)^2\}$. Thus the general formula for the significance levels of z derivable from (1) is

$$z = \frac{x}{\sqrt{h-1}} - \left(\frac{x^2+2}{6}\right)\left(\frac{1}{n_1} - \frac{1}{n_2}\right),$$

where x is a normal deviate with unit standard error. By inserting the appropriate significance level of x, this formula has been extended [2] to the tables of the 20, 1, and 0.1 percent levels of z and commonly appears with all published tables of z. The objects of this note are to indicate the derivation of the formula and to suggest an improvement upon it in the latter cases.

2. **The transformation of the z-distribution to normality.** For high values of n_1 and n_2, the distribution of z approaches the normal distribution, the principal deviation being a slight skewness introduced by the inequality of n_1 and n_2. It is therefore natural to seek an approximate formula for the distribution of z by examining its relation to the normal distribution. For the z-distribution the ratio $\kappa_r/\kappa_2^{r/2}$, where κ_r is the r^{th} cumulant, is of the order $n^{-(\frac{1}{2}r-1)}$, where n is the smaller of n_1 and n_2. This property is common to a large number of distributions which tend to normality; for example, the distribution of the mean of a sample of size n from any distribution with finite cumulants. Fisher and Cornish [3] have recently given a method, applicable to all distributions with this property, for transforming the distribution to a normal distribution to any desired order of approximation. They also obtained explicit expressions for the significance levels of the original distribution in terms of the significance levels of the normal distribution, discussing the z-distribution as a particular example. The relation between z and the normal deviate x at the same level of probability was found to be

$$(2) \quad z = \frac{x}{\sqrt{h}} - \tfrac{1}{6}(x^2+2)\left(\frac{1}{n_1} - \frac{1}{n_2}\right) + \frac{1}{\sqrt{h}}\left\{\frac{x^3+3x}{12\,h} + \frac{x^3+11x}{144}h\left(\frac{1}{n_1} - \frac{1}{n_2}\right)^2\right\},$$

the three terms on the right hand side being respectively of order $n^{-\frac{1}{2}}$, n^{-1}, and $n^{-\frac{3}{2}}$, so that terms of order n^{-2} are neglected.[1]

If this equation is compared with equation (1), the latter appears at first sight to be the approximation of order n^{-1} to the z-distribution, except that the divisor of x is $\sqrt{h-1}$ in (1) and $\sqrt{h}$ in (2). Computation of a few values shows that at the 5 percent level, equation (1) is the better approximation. For example, for $n_1 = 40$, $n_2 = 60$, (1) gives z (5 percent) $= .2334$, (2) gives $.2309$, and the exact value is $.2332$.

Since

$$\frac{x}{\sqrt{h-1}} = \frac{x}{\sqrt{h}} + \frac{x}{2h\sqrt{h}} + \text{terms of order } n^{-2},$$

Fisher's approximation differs from (2) by including a correction term of order $n^{-\frac{3}{2}}$. Inspection of the true correction terms of this order in equation (2) shows that for finite values of n_1 and n_2 the term $\dfrac{x^3+11x}{144}\sqrt{h}\left(\dfrac{1}{n_1} - \dfrac{1}{n_2}\right)^2$ is considerably smaller than the term $\dfrac{x^3+3x}{12h\sqrt{h}}$, since the former has a smaller numerical coefficient and involves the difference between $\dfrac{1}{n_1}$ and $\dfrac{1}{n_2}$. Thus Fisher's formula gives a close approximation to the true formula of order $n^{-\frac{3}{2}}$, provided that $\dfrac{x}{2}$ is approximately equal to $\dfrac{x^3+3x}{12}$; i.e. if $\dfrac{x^2+3}{6}$ is approximately equal

[1] Fisher and Cornish also gave the two succeeding terms.

to 1. For the 5 percent level, $x = 1.6449$, and $\frac{x^2 + 3}{6} = 0.951$. Thus at the 5 percent level the use of $\sqrt{h - 1}$ in (1) instead of $\sqrt{h}$ extends the validity of Fisher's approximation from order n^{-1} to order $n^{-\frac{3}{2}}$.

This ingenious device, however, requires adjustment at other levels of significance. The values of $(x^2 + 3)/6$ at the principal significance levels are shown below.

Significance level—%	40	30	20	10	5	1	0.1
$\lambda = (x^2 + 3)/6$	0.51	0.55	0.62	0.77	0.95	1.40	2.09

If $\sqrt{h - 1}$ in formula (1) is replaced by $\sqrt{h - \lambda}$, with the above values of λ, Fisher's formula will be approximately valid to order $n^{-\frac{3}{2}}$ at all levels of significance. In particular, for the tables already published of the 20, 1 and 0.1 percent points, λ may be taken as 0.6, 1.4 and 2.1 respectively. The values of z given by the use of $\sqrt{h - 1}$ and $\sqrt{h - \lambda}$ are compared below for $n_1 = 24$, $n_2 = 60$.[2]

Significance Level	Approximate formula		Exact value
	$\sqrt{h - 1}$	$\sqrt{h - \lambda}$	
20%	.1346	.1337	.1338
1%	.3723	.3748	.3746
0.1%	.4875	.4966	.4955

The use of $\sqrt{h - \lambda}$ gives values practically correct to 4 decimal places, except for the 0.1 level of significance, at which the higher terms become more important.

With the aid of this formula, complete tabulation of the z-distribution for a given pair of high values of n_1 and n_2 is relatively simple. If very low probabilities at the tails are required, the further approximations given by Fisher and Cornish [3] may be used.

REFERENCES

[1] R. A. Fisher. *Statistical Methods for Research Workers.* Edinburgh, Oliver and Boyd. 1st Ed. 1925.

[2] R. A. Fisher and F. Yates. *Statistical Tables.* Edinburgh, Oliver and Boyd. 1938.

[3] E. A. Cornish and R. A. Fisher. "Moments and Cumulants in the Specification of Distributions," *Revue de l'Institut International de Statistique*, Vol. 4 (1937).

Iowa State College,
Ames, Iowa.

[2] The numerical terms in the approximate formula given for the 20 percent points on p. 28 of Fisher and Yates' *Statistical Tables* are in error. Their formula should read:

$$z = \frac{0.8416}{\sqrt{h - 1}} - 0.4514\left(\frac{1}{n_1} - \frac{1}{n_2}\right)$$

A NOTE ON THE ANALYSIS OF VARIANCE WITH UNEQUAL CLASS FREQUENCIES[1]

B. ABRAHAM WALD[2]

Let us consider p groups of variates and denote by m_j $(j = 1, \cdots, p)$ the number of elements in the j-th group. Let x_{ij} be the i-th element of the j-th group. We assume that x_{ij} is the sum of two variates ϵ_{ij} and η_j, i.e. $x_{ij} = \epsilon_{ij} + \eta_j$, where ϵ_{ij} $(i = 1, \cdots, m_j; j = 1, \cdots, p)$ is normally distributed with mean μ and variance σ^2, and η_j $(j = 1, \cdots, p)$ is normally distributed with mean μ' and variance σ'^2. All the variates ϵ_{ij} and η_j are supposed to be distributed independently.

The intraclass correlation ρ is given by[3]

$$\rho = \frac{\sigma'^2}{\sigma^2 + \sigma'^2}.$$

Confidence limits for ρ have been derived only in case of equal class frequencies, i.e. $m_1 = m_2 = \cdots = m_p$. In this paper we shall deal with the problem of determining the confidence limits for ρ in the case of unequal class frequencies. Since ρ is a monotonic function of $\frac{\sigma'^2}{\sigma^2}$, our problem is solved if we derive confidence limits for $\frac{\sigma'^2}{\sigma^2}$.

Denote by $\bar{x}_j$ the arithmetic mean of the j-th group, i.e.

$$\bar{x}_j = \frac{\sum_{i=1}^{m_j} \epsilon_{ij}}{m_j} + \eta_j. \tag{1}$$

Hence the variance of $\bar{x}_j$ is equal to

$$\sigma^2_{\bar{x}_j} = \frac{\sigma^2}{m_j} + \sigma'^2. \tag{2}$$

Denote $\frac{\sigma'^2}{\sigma^2}$ by λ^2. Then we have

$$\sigma^2_{\bar{x}_j} = \sigma^2\left(\frac{1}{m_j} + \lambda^2\right) = \frac{\sigma^2}{w_j}, \tag{3}$$

---

[1] The author is indebted to Professor H. Hotelling for formulating the problem dealt with in this paper.

[2] Research under a grant-in-aid from the Carnegie Corporation at New York.

[3] See for instance R. A. Fisher, *Statistical Methods for Research Workers*, 6-th edition, p. 228.

where

(4) $$w_j = \frac{m_j}{1 + m_j\lambda^2}.$$

Now we shall prove that

(5) $$\frac{1}{\sigma^2}\sum_{j=1}^{p}\left[w_j\left(\bar{x}_j - \frac{\sum_{j=1}^{p} w_j\bar{x}_j}{\sum_{j=1}^{p} w_j}\right)^2\right]$$

has the χ^2-distribution with $p - 1$ degrees of freedom. Let

$$y_j = \sqrt{w_j}\,\bar{x}_j \qquad (j = 1, \cdots, p)$$

and consider the orthogonal transformation

$$y_1' = L_1(y_1, \cdots, y_p),$$

$$\cdots\cdots\cdots\cdots$$

$$y_{n-1}' = L_{p-1}(y_1, \cdots, y_p),$$

$$y_p' = L_p(y_1, \cdots, y_p) = \frac{\sqrt{w_1}\,y_1 + \cdots + \sqrt{w_p}\,y_p}{\sqrt{w_1 + \cdots + w_p}},$$

where $L_1(y_1, \cdots, y_p), \cdots, L_{p-1}(y_1, \cdots, y_p)$ denote arbitrary homogenous linear functions subject to the only condition that the transformation should be orthogonal.

Since the mean value of y_j is equal to $\sqrt{w_j}\,(\mu + \mu')$ and the variance of y_j is equal to σ^2, we obviously have: The mean value of y_j' $(j = 1, \cdots, p - 1)$ is equal to zero, the variance of y_j' $(j = 1, \cdots, p)$ is equal to σ^2. In order to prove our statement, we have only to show that the expression (5) is equal to $\frac{1}{\sigma^2}(y_1'^2 + \cdots + y_{p-1}'^2)$. If we substitute in (5) $\frac{y_j}{\sqrt{w_j}}$ for $\bar{x}_j$, we get

$$\frac{1}{\sigma^2}\sum_{j=1}^{p}\left\{w_j\left[\frac{y_j^2}{w_j} - 2\,\frac{y_j}{\sqrt{w_j}}\,\frac{\sum_{j=1}^{p}\sqrt{w_j}\,y_j}{\sum_{j=1}^{p} w_j} + \left(\frac{\sum_j \sqrt{w_j}\,y_j}{\Sigma w_j}\right)^2\right]\right\}$$

(5') $$= \frac{1}{\sigma^2}\left[\sum_j y_j^2 - 2\,\frac{(\sum_j \sqrt{w_j}\,y_j)^2}{\sum_j w_j} + \frac{(\sum_j \sqrt{w_j}\,y_j)^2}{\sum_j w_j}\right]$$

$$= \frac{1}{\sigma^2}\left[\sum_j y_j^2 - \frac{(\sum_j \sqrt{w_j}\,y_j)^2}{\sum_j w_j}\right] = \frac{1}{\sigma^2}\left[\sum_{j=1}^{p} y_j^2 - y_p'^2\right] = \frac{1}{\sigma^2}\left[\sum_{j=1}^{p} y_j'^2 - y_p'^2\right]$$

$$= \frac{1}{\sigma^2}(y_1'^2 + \cdots + y_{p-1}'^2).$$

Since $\frac{\Sigma\Sigma(x_{ij} - \bar{x}_j)^2}{\sigma^2}$ has the χ^2 distribution with $N - p$ degrees of freedom, the expression

(6) $$F = \frac{N-p}{p-1} \frac{\sum_{j=1}^{p} \left\{ w_j \left(\bar{x}_j - \frac{\Sigma w_j \bar{x}_j}{\Sigma w_j} \right)^2 \right\}}{\Sigma\Sigma(x_{ij} - \bar{x}_j)^2}$$

has the analysis of variance distribution with $p - 1$ and $N - p$ degrees of freedom, where $N = m_1 + \cdots + m_p$. In case $m_1 = m_2 = \cdots = m_p = m$, we have

(6′) $$F = \frac{N-p}{p-1} \frac{\sum_{j=1}^{p} (\bar{x}_j - \bar{x})^2}{\Sigma\Sigma(x_{ij} - \bar{x}_j)^2} \cdot \frac{m}{1 + m\lambda^2} = \frac{1}{1 + m\lambda^2} F^*,$$

where $\bar{x} = \frac{\Sigma\Sigma x_{ij}}{N}$ and $F^* = \frac{N-p}{p-1} \frac{m\Sigma(\bar{x}_j - \bar{x})^2}{\Sigma\Sigma(x_{ij} - \bar{x}_j)^2}$.

Hence

$$\lambda^2 = \left(\frac{F^*}{F} - 1 \right) \frac{1}{m}.$$

If F_1 denotes the lower and F_2 the upper confidence limit of F, we obtain for λ^2 the confidence limits

$$\left(\frac{F^*}{F_1} - 1 \right) \frac{1}{m} \quad \text{and} \quad \left(\frac{F^*}{F_2} - 1 \right) \frac{1}{m}.$$

Let us now consider the general case that $m_1, \cdots, m_p$ are arbitrary positive integers. First we shall show that the set of values of λ^2, for which (6) lies between its confidence limits F_1 and F_2, is an interval. For this purpose we have only to show that

$$f(\lambda^2) \equiv \sum_{j=1}^{p} \left\{ w_j \left(\bar{x}_j - \frac{\Sigma w_j \bar{x}_j}{\Sigma w_j} \right)^2 \right\}$$

is monotonically decreasing with λ^2. In fact

$$\frac{df(\lambda^2)}{d\lambda^2} = \sum_{j=1}^{p} \frac{dw_j}{d\lambda^2} \left(\bar{x}_j - \frac{\Sigma w_j \bar{x}_j}{\Sigma w_j} \right)^2 - 2 \frac{d}{d\lambda^2} \left(\frac{\Sigma w_j \bar{x}_j}{\Sigma w_j} \right) \left[\sum_{j=1}^{p} w_j \left(\bar{x}_j - \frac{\Sigma w_j \bar{x}_j}{\Sigma w_j} \right) \right].$$

Since

$$\sum_{j=1}^{p} w_j \left(\bar{x}_j - \frac{\Sigma w_j \bar{x}_j}{\Sigma w_j} \right) = 0,$$

we have

$$\frac{df(\lambda^2)}{d\lambda^2} = \sum_{j=1}^{p} \frac{dw_j}{d\lambda^2} \left(\bar{x}_j - \frac{\Sigma w_j \bar{x}_j}{\Sigma w_j} \right)^2 = \sum_{j=1}^{p} - w_j^2 \left(\bar{x}_j - \frac{\Sigma w_j \bar{x}_j}{\Sigma w_j} \right)^2 < 0,$$

which proves our statement.

Hence the lower confidence limit λ_1^2 of λ^2 is given by the root of the equation in λ^2:

$$(7) \qquad F = \frac{N-p}{p-1} \frac{\sum_{j=1}^{p} \left\{ w_j \left(\bar{x}_j - \frac{\Sigma w_j \bar{x}_j}{\Sigma w_j} \right)^2 \right\}}{\Sigma\Sigma (x_{ij} - \bar{x}_j)^2} = F_2$$

and the upper confidence limit λ_2^2 of λ^2 is given by the root of the equation in λ^2:

$$(8) \qquad F = F_1 .$$

Since $f(\lambda^2)$ is monotonically decreasing, the equations (7) and (8) have at most one root in λ^2. If the equation (7) or (8) has no root, the corresponding confidence limit has to be put equal to zero. If neither (7) nor (8) has a root, we have to reject at least one of the hypotheses:

(1) $x_{ij} = \epsilon_{ij} + \eta_j$.

(2) The variates ϵ_{ij} and η_j ($i = 1, \cdots, m_j$; $j = 1, \cdots, p$) are normally and independently distributed.

(3) Each of the variates ϵ_{ij} has the same distribution.

(4) Each of the variates η_j has the same distribution.

The equations (7) and (8) are complicated algebraic equations in λ^2. For the actual calculation of the roots of these equations, well known approximation methods can be applied making use also of the fact that the left members are monotonic functions of λ^2. In applying any approximation method it is very useful to start with two limits of the root which do not lie far apart. We shall give here a method of finding such limits.

Denote by $\bar{F}$ the function which we obtain from F (formula (6)) by substituting

$$\bar{w}_j = \frac{l_j}{1 + l_j \lambda^2} \text{ for } w_j \qquad (j = 1, \cdots, p).$$

Let $\bar{f}$ be the function obtained from f by the same process.

Denote by $\varphi(m, \lambda^2)$ the function which we obtain from $\bar{F}$ by substituting m for $l_1, \cdots, l_p$. We shall first show that $\bar{F}$ is non-decreasing with increasing l_k ($k = 1, \cdots, p$), i.e. $\frac{\partial \bar{F}}{\partial l_k} \geq 0$. For this purpose we have only to show that $\frac{\partial \bar{f}}{\partial l_k} \geq 0$. We have:

$$\frac{\partial \bar{f}}{\partial l_k} = \sum_j \frac{\partial \bar{w}_j}{\partial l_k} \left(\bar{x}_j - \frac{\Sigma \bar{w}_j \bar{x}_j}{\Sigma \bar{w}_j} \right)^2 - 2 \frac{\partial}{\partial l_k} \left(\frac{\Sigma \bar{w}_j \bar{x}_j}{\Sigma \bar{w}_j} \right) \cdot \left[\Sigma \bar{w}_j \cdot \left(\bar{x}_j - \frac{\Sigma \bar{w}_j \bar{x}_j}{\Sigma \bar{w}_j} \right) \right]$$

$$= \sum_j \frac{\partial \bar{w}_j}{\partial l_k} \left(\bar{x}_j - \frac{\Sigma \bar{w}_j \bar{x}_j}{\Sigma \bar{w}_j} \right)^2 = \frac{1}{(1 + l_k \lambda^2)^2} \left(\bar{x}_k - \frac{\Sigma \bar{w}_j \bar{x}_j}{\Sigma \bar{w}_j} \right)^2 \geq 0.$$

Hence our statement is proved. Denote by m' the smallest and by m'' the greatest of the values $m_1, \cdots, m_p$. Then we obviously have

$$\varphi(m', \lambda^2) \leq F \leq \varphi(m'', \lambda^2). \tag{9}$$

Denote by $\lambda_1'^2$, $\lambda_1''^2$, $\lambda_2'^2$, $\lambda_2''^2$ the roots in λ^2 of the following equations respectively:

$$\varphi(m', \lambda^2) = F_2 ;$$

$$\varphi(m'', \lambda^2) = F_2 ;$$

$$\varphi(m', \lambda^2) = F_1 ; \qquad \varphi(m'', \lambda^2) = F_1 .$$

Since F is monotonically decreasing with increasing λ^2, on account of (7), (8), and (9) we obviously have

$$\lambda_1'^2 \leq \lambda_1^2 \leq \lambda_1''^2$$

and

$$\lambda_2'^2 \leq \lambda_2^2 \leq \lambda_2''^2.$$

The above inequalities give us the required limits.

COLUMBIA UNIVERSITY,
NEW YORK, N. Y.

THE DISTRIBUTION OF QUADRATIC FORMS IN NON-CENTRAL NORMAL RANDOM VARIABLES

BY WILLIAM G. MADOW[1]

The following theorem is the algebraic basis of the theorem of R. A. Fisher and W. G. Cochran which states necessary and sufficient conditions that a set of quadratic forms in normally and independently distributed random variables should themselves be independently distributed in χ^2-distributions.[2]

THEOREM I. *If the real quadratic forms* $q_1, \cdots, q_m$, *in* $x_1, \cdots, x_n$, *are such that*

$$\sum_\gamma q_\gamma = \sum_\nu x_\nu^2, \tag{1}$$

and if the rank of q_γ *is* n_γ, *then a necessary and sufficient condition that*

$$q_\gamma = \sum_\alpha z_\alpha^2, \tag{2}$$

[1] The letters i, j, μ, ν will assume all integral values from 1 through n, the letter γ will assume all integral values from 1 through m, $(n \geq m)$, the letter α will assume all integral values from $n_1 + \cdots + n_{\gamma-1} + 1$ through $n_1 + \cdots + n_\gamma$, $(n_0 = 0, n_1 + \cdots + n_m = n')$, the letters β, β' will assume all integral values from 1 through n', and the letters r, s will assume all integral values from 1 through $n - 1$.

[2] The references are: W. G. Cochran, "The Distribution of Quadratic Forms in a Normal System, with Applications to the Analysis of Covariance," *Proc. Camb. Phil. Soc.*, Vol. 30 (1934), pp. 178–191, and R. A. Fisher, "Applications of 'Student's' Distribution," *Metron*, Vol. 5 (1926), pp. 90–104.

where the real linear functions z_β of the x_ν are defined by

$$x_\nu = \sum_\beta c_{\nu\beta} z_\beta \tag{3}$$

is

$$n' = n. \tag{4}$$

Furthermore the system of linear forms (3) *constitute an orthogonal transformation.*

PROOF: *Necessity.* Since the rank of a sum of quadratic forms is less than or equal to the sum of their ranks, it follows that $n' \geq n$. Upon substituting from (3) for the x's in (1), and using (2), it is seen that, for all values of the z's,

$$\sum_\beta z_\beta^2 = \sum_{\beta,\beta'} \left(\sum_\nu c_{\nu\beta} c_{\nu\beta'}\right) z_\beta z_{\beta'}$$

and hence, from (1), it follows that

$$\sum_\nu c_{\nu\beta} c_{\nu\beta'} = \delta_{\beta\beta'} \tag{5}$$

where $\delta_{\beta\beta'} = 0$, if $\beta \neq \beta'$, and $\delta_{\beta\beta'} = 1$ if $\beta = \beta'$. However, since the rank of the system of linear forms (3) is not greater than n, and since the matrix of (5) is the product of the matrix of (3) by its transposed matrix, it follows that (5) can be true only if n' is not greater than n. Consequently $n' = n$. It then is an immediate result of (5) that the transformation (3) is orthogonal.

Sufficiency. We assume that $n' = n$. By a real linear transformation of $x_1, \cdots, x_n$ we obtain linear forms z_ν such that

$$q_\gamma = \sum_\alpha c_\alpha z_\alpha^2,$$

where $c_\alpha = 1$ or -1. The set of linear functions $z_1, \cdots, z_n$ are linearly independent, for if $z_n \not\equiv 0$, and if real numbers $h_1, \cdots, h_{n-1}$ not all zero, exist such that, say,

$$z_n = \sum_r h_r z_r$$

then

$$\sum_\nu z_\nu^2 = \sum_{r,s} H_s z_r z_s.$$

Substituting, we have

$$\sum_\gamma q_\gamma = \sum_\nu c_\nu z_\nu^2 = \sum_{r,s} \sum_{\mu,\nu} H_{rs} c^{r\mu} c^{s\nu} x_\mu x_\nu$$

where $z_\nu = \sum_\mu c^{\nu\mu} x_\mu$. (It is not assumed here that the matrix of the $c^{\mu\nu}$ is the inverse of the matrix of the $c_{\mu\nu}$. That fact is a consequence of this proof.)

Denoting the matrix of $z_1, \cdots, z_{n-1}$ by $\bar{C}_n$ we see that the matrix of $\sum_\gamma q_\gamma$ is $\bar{C}'_m H \bar{C}_n$ where H is the matrix of the H_{rs} and has rank less than or equal to $n - 1$ which contradicts the hypothesis. Hence if C is the matrix having the elements

c_ν in its main diagonal and zeros elsewhere and if C_n is the matrix of $z_1, \cdots, z_n$ it follows that

$$C_n' C C_n = I,$$

where I is the identity matrix, i.e. the matrix having ones in the main diagonal and zeros elsewhere and C_n non-singular. Then $C = C_n^{-1\prime} C_n^{-1}$ and hence C is the identity matrix and C_n is orthogonal.

Among the hypotheses of the Fisher-Cochran theorem is the hypothesis that the mean value of x_μ is 0, and the variance of x_μ is σ^2. However, in connection with his analysis of the distribution of the multiple correlation coefficient,[3] R. A. Fisher derived the distribution of the sum of the squares of n independently distributed random variables $x_1, \cdots, x_n$, the probability density of x_μ being given by

$$p(x_\mu) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left[-\frac{1}{2\sigma^2}(x_\mu - a_\mu)^2\right]. \tag{6}$$

More recently, P. C. Tang,[4] has used the distribution of the sum of non-central squares in his study of the power function of the analysis of variance test.

In this note we extend the Fisher-Cochran theorem to non-central random variables. If the random variables x_μ are independently distributed with probability densities given by (6), Fisher and Tang have shown that if $\chi'^2 = \frac{1}{\sigma^2} \sum_\nu x_\nu^2$, then the probability density of χ'^2 is given by

$$p(\chi'^2) = \tfrac{1}{2} e^{-\lambda} (\tfrac{1}{2}\chi'^2)^{\frac{1}{2}n-1} e^{-\frac{1}{2}\chi'^2} \sum_{\nu=0}^{\infty} \frac{(\frac{1}{2}\lambda\chi'^2)^\nu}{\nu!\,\Gamma(\frac{1}{2}n + \nu)}, \tag{7}$$

where $\lambda = \frac{1}{2\sigma^2} \sum_\nu a_\nu^2$.

We now give necessary and sufficient conditions that a set of quadratic forms in normally and independently distributed random variables should themselves be independently distributed in χ'^2-distributions.

THEOREM II. *Let $x_1, \cdots, x_n$ be independently distributed random variables, the random variable x_μ having probability density* (6). *Denote $\sum_\nu x_\nu^2$ by q, and denote $\frac{1}{2\sigma^2} \sum_\nu a_\nu^2$ by λ. Let $q_1, \cdots, q_m$, be quadratic forms,*

$$q_\gamma = \sum_{\mu,\nu} a_{\mu\nu}^{(\gamma)} x_\mu x_\nu$$

such that $\sum_\gamma q_\gamma = q$, and let the rank of q_γ be denoted by n_γ.

---

[3] R. A. Fisher, "The General Sampling Distribution of the Multiple Correlation Coefficient," *Proc. Royal Soc. of London,* (A), Vol. 121 (1928), pp. 654–673.

[4] P. C. Tang, "The Power Function of the Analysis of Variance Tests with Tables and Illustrations of their Use," *Statistical Research Memoirs,* Vol. 2 (1938), pp. 126–149.

A necessary and sufficient condition that the quadratic forms $\chi_\gamma'^2$, $\left(\chi_\gamma'^2 = \frac{q_\gamma}{\sigma^2}\right)$, *be independently distributed with joint probability density*

$$p(\chi_1'^2, \cdots, \chi_m'^2) = \prod_\gamma p(\chi_\gamma'^2), \tag{8}$$

where $p(\chi_\gamma'^2)$ *is given by* (7) *with* n_γ *and* λ_γ *in place of* n *and* λ, *and*

$$\lambda_\gamma = \frac{1}{2\sigma^2} \sum_{\mu,\nu} a_{\mu\nu}^{(\gamma)} a_\mu a_\nu \tag{9}$$

is $n' = n$.

Proof. *Necessity.* Tang[5] has shown that the distribution of χ'^2 is given by (7) and that if the $\chi_\gamma'^2$ have joint distribution (8), then the distribution of $\chi_1'^2 + \cdots + \chi_m'^2$, $(= \chi'^2)$, is (7) with n' in place of n. Upon comparing terms, we see that $n' = n$.

Sufficiency. By Theorem I there exist n orthogonal linear functions (3) such that (2) is true. Then it is easy to see that the random variables $z_1, \cdots, z_n$ are independently distributed with a joint probability density

$$p(z_1, \cdots, z_n) = (2\pi\sigma^2)^{-\frac{1}{2}n} \exp\left[-\tfrac{1}{2}\sum_\nu (z_\nu - a_\nu')^2\right], \tag{10}$$

where

$$\sum_\nu a_\nu'^2 = \sum_\nu a_\nu^2, \quad \text{and} \quad a_\mu' = \sum_\nu c_{\mu\nu} a_\nu .$$

If we set $2\sigma^2\lambda_\gamma = \sum_\alpha a_\alpha'^2$, then we have, from (7) and (10), that the $\chi_\gamma'^2$ are independently distributed with joint probability density (8). It is only necessary to show that $\sum_\alpha a_\alpha'^2 = \sum_{\mu,\nu} a_{\mu\nu}^{(\gamma)} a_\mu a_\nu$ in order to complete the proof of the theorem. Now

$$\sum_{\mu,\nu} a_{\mu\nu}^{(\gamma)} a_\mu a_\nu = \sum_{i,j} \left(\sum_{\mu,\nu} a_{\mu\nu}^{(\gamma)} c_{i\mu} c_{j\nu}\right) a_i' a_j' .$$

On the other hand, by direct substitution for the z's we see that

$$q_\gamma = \sum_\alpha z_\alpha^2 = \sum_{\mu,\nu} \left(\sum_\alpha c_{\mu\alpha} c_{\nu\alpha}\right) x_\mu x_\nu$$

and hence $a_{\mu\nu}^{(\gamma)} = \sum_\alpha c_{\mu\alpha} c_{\nu\alpha}$. Since (1) is an orthogonal transformation,

$$\sum_{\mu,\nu} a_{\mu\nu}^{(\gamma)} c_{i\mu} c_{j\nu} = \sum_{\mu,\nu} \left(\sum_\alpha c_{\mu\alpha} c_{\nu\alpha}\right) c_{i\mu} c_{j\nu} = \sum_\alpha \delta_{\alpha i} \delta_{\alpha j} ,$$

where $\delta_{\alpha i} = 0$, if $\alpha \neq i$ and $= 1$ if $\alpha = i$, which completes the proof.

It is emphasized that the form of λ_γ makes it unnecessary to calculate the matrix of q_γ to determine λ_γ since the values a_ν need only be substituted for the x_ν in the original expression for q_γ to determine λ_γ.

Washington, D. C.

---

[5] See 4 p. 140.

TWO PROPERTIES OF SUFFICIENT STATISTICS

By Louis Olshevsky

The concept of sufficient statistics was introduced by R. A. Fisher in 1922. It was refined and extended in 1936 by Neyman and Pearson who gave definitions of shared sufficient statistics and sufficient sets of algebraically independent statistics.[1] Today the concept plays an important part in the theory of the subject. Characterized briefly, a statistic associated with a single or specific population parameter is sufficient when no other statistic calculated from the same sample sheds any additional light on the value of the parameter. We shall prove that sets of sufficient statistics possess certain interconnections so that when one set is known every other set with a like number of members and linked with the same population parameters is discoverable.

Theorem 1. *If $T_1, \cdots, T_m$ are a set of m ($m \leqq n$) algebraically independent sufficient statistics with regard to the parameters $\theta_1, \cdots, \theta_q$ and the probability law $p(x_1, \cdots, x_n \mid \theta_1, \cdots, \theta_q, \cdots, \theta_l)$, a necessary and sufficient condition for the sufficiency of any set of m algebraically independent statistics $T'_1, \cdots, T'_m$ with regard to the same parameters and the same probability distribution is that the T'_i be a set of independent functions of the T_j ($i, j = 1, \cdots, m$).*

Proof: As an adjunct in the demonstration we cite the following theorem due to Neyman.[2] For a set of algebraically independent statistics $T_1, \cdots, T_m$ to be a sufficient set with regard to the parameters $\theta_1, \cdots, \theta_q$, it is necessary and sufficient that in any point of sample space, except perhaps for a set of measure zero, it should be possible to present the probability law in the form of the product

$$\begin{aligned} &p(x_1, \cdots, x_n \mid \theta_1, \cdots, \theta_q, \cdots, \theta_l) \\ (1) \qquad &\qquad = p(T_1, \cdots, T_m \mid \theta_1, \cdots, \theta_q)\cdot\phi(x_1, \cdots, x_n ; \theta_{q+1}, \cdots, \theta_l) \end{aligned}$$

where $p(T_1, \cdots, T_m \mid \theta_1, \cdots, \theta_q)$ is the probability law of $T_1, \cdots, T_m$ and the function ϕ does not depend upon $\theta_1, \cdots, \theta_q$.

The sufficiency of the condition stated in the hypothesis of Theorem I is now immediately evident. For, if p' and ϕ' refer to the second set of algebraically independent statistics and $T'_i = T'_i(T_1, \cdots, T_m)$ where the functions are independent, the relations can be solved for the T_j in terms of the T'_i giving

$$T_j = T_j(T'_1, \cdots, T'_m), \quad p'(T'_1, \cdots, T'_m \mid \theta_1, \cdots, \theta_q)$$
$$= p[T_1(T'_1, \cdots, T'_m), \cdots, T_m(T'_1, \cdots, T'_m) \mid \theta_1, \cdots, \theta_q] \frac{\partial(T_1, \cdots, T_m)}{\partial(T'_1, \cdots, T'_m)},$$
$$\phi'(x_1, \cdots, x_n; \theta_{q+1}, \cdots, \theta_l) = \phi(x_1, \cdots, x_n; \theta_{q+1}, \cdots, \theta_l) \div \frac{\partial(T_1, \cdots, T_m)}{\partial(T'_1, \cdots, T'_m)},$$

---

[1] See Neyman and Pearson: "Sufficient Statistics and Uniformly Most Powerful Tests of Statistical Hypotheses," *Statistical Research Memoirs of the University of London*, June 1936. The notatian of the present paper is taken from this article.

[2] See Neyman's article in the *Giornale dell' Insituto Italiano degli Attuari*, Vol. VI, No. 4 (1935) as well as the memoir referred to in footnote 1.

and

$$(2)\quad \begin{aligned} &p(x_1, \cdots, x_n \mid \theta_1, \cdots, \theta_q, \cdots, \theta_l) \\ &\qquad = p'(T'_1, \cdots, T'_m \mid \theta_1, \cdots, \theta_q) \cdot \phi'(x_1, \cdots, x_n; \theta_{q+1}, \cdots, \theta_l). \end{aligned}$$

Proof of the necessity is somewhat more involved. Since the T_j and T'_i are both sets of algebraically independent statistics with regard to $\theta_1, \cdots, \theta_q$, equations (1) and (2) are satisfied. They are, in fact, identities when the values of $T_1, \cdots, T_m$ and $T'_1, \cdots, T'_m$ in terms of the x_i are substituted. Division of (1) by (2) and multiplication leads to the equation

$$(3)\quad \frac{p(T_1, \cdots, T_m \mid \theta_1, \cdots, \theta_q)}{p'(T'_1, \cdots, T'_m \mid \theta_1, \cdots, \theta_q)} = \frac{\phi'(x_1, \cdots, x_n; \theta_{q+1}, \cdots, \theta_l)}{\phi(x_1, \cdots, x_n; \theta_{q+1}, \cdots, \theta_l)}.$$

The right side of (3) is free of $\theta_1, \cdots, \theta_q$. Therefore, in reality the left side must be too. If some or all of the parameters $\theta_1, \cdots, \theta_q$ enter formally into the left side, we can choose $m + 1$ sets of values $\theta_1^i, \cdots, \theta_q^i$ $(i = 1, \cdots, m + 1)$ such that each of the $m + 1$ functions $p(T_1, \cdots, T_m \mid \theta_1^i, \cdots, \theta_q^i) \div p'(T'_1, \cdots, T'_m \mid \theta_1^i, \cdots, \theta_q^i)$ differs formally from all of the others. We can, then, since each is equal to the right side of (3) which is free of $\theta_1, \cdots, \theta_q$, equate any one of these functions to the remaining m in turn. This provides m independent equations whose very existence proves that the T'_i are functions of the T_j and vice versa.

If none of the parameters $\theta_1, \cdots, \theta_q$ enters formally into the left side of (3), $p(T_1, \cdots, T_m \mid \theta_1, \cdots, \theta_q)$ must be of the form $p(T_1, \cdots, T_m)g(\theta_1, \cdots, \theta_q)$ and $p'(T'_1, \cdots, T'_m \mid \theta_1, \cdots, \theta_q)$ of the form $p'(T'_1, \cdots, T'_m)g(\theta_1, \cdots, \theta_q)$. In this case the original probability law $p(x_1, \cdots, x_n \mid \theta_1, \cdots, \theta_q, \cdots, \theta_l)$ contains $\theta_1, \cdots, \theta_q$ only nominally and there can be no talk of any statistics designed to estimate these parameters either singly or in combination.

When $m = 1$ and the set of algebraically independent statistics reduces to one, the single statistic is termed a shared sufficient statistic of the parameters $\theta_1, \cdots, \theta_q$.[3] For this special case, Theorem I can be restated as follows. If T is a shared sufficient statistic with regard to the population parameters $\theta_1, \cdots, \theta_q$ and the probability distribution $p(x_1, \cdots, x_n \mid \theta_1, \cdots, \theta_q, \cdots, \theta_l)$, the necessary and sufficient condition for the sufficiency of any statistic T' with regard to the same parameters and the same probability distribution is that T' be a function of T. When m and q both equal one, the statistic becomes a sufficient statistic in the sense originally defined by Fisher in 1922.

A physical law is independent of the coordinate system used to express it. This fact is taken account of in modern physics through the employment of tensors. One might hope for a parallel situation in the relation between sufficient statistics and the probability law to which they refer. Given any l parameter family of distribution laws $p(x_1, \cdots, x_n \mid \theta_1, \cdots, \theta_l)$, the substitu-

---

[3] See the memoir mentioned in footnote 1.

tion $\theta_i = \theta_i(\theta_1', \cdots, \theta_l')$ $(i = 1, \cdots, l)$ leads to the equally valid representation of the family

$$p'(x_1, \cdots, x_n \mid \theta_1', \cdots, \theta_l') = p[x_1, \cdots, x_n \mid \theta_1(\theta_1', \cdots, \theta_l'), \cdots, \theta_l(\theta_1', \cdots, \theta_l')].$$

Is a set of statistics sufficient with respect to the first representation also sufficient with respect to the second? The answer is partly in the affirmative and is given by the following proposition.

THEOREM II. *If the set of algebraically independent statistics $T_1, \cdots, T_m$ is sufficient with regard to the parameters $\theta_1, \cdots, \theta_q$ and the probability law $p(x_1, \cdots, x_n \mid \theta_1, \cdots, \theta_q, \cdots, \theta_l)$, it is also sufficient with regard to $\theta_1', \cdots, \theta_q'$ and any other representation $p'(x_1, \cdots, x_n \mid \theta_1', \cdots, \theta_q', \cdots, \theta_l')$ of the same probability law provided θ_i' $(i = 1, \cdots, q)$ are independent functions of $\theta_1, \cdots, \theta_q$ only and θ_j' $(j = q + 1, \cdots, l)$ are functions of $\theta_{q+1}, \cdots, \theta_l$ only.*

PROOF: The proof of the theorem is obvious. We are given the fact that $p(x_1, \cdots, x_n \mid \theta_1, \cdots, \theta_q, \cdots, \theta_l) = p(T_1, \cdots, T_m \mid \theta_1, \cdots, \theta_q)\cdot\phi(x_1, \cdots, x_n ; \theta_{q+1}, \cdots, \theta_l)$. Since the θ_i' $(i = 1, \cdots, q)$ are functions of $\theta_1, \cdots, \theta_q$ only and the θ_j' $(j = q + 1, \cdots, l)$ are functions of $\theta_{q+1}, \cdots, \theta_l$ only, it follows that $\theta_i = \theta_i(\theta_1', \cdots, \theta_q')$ $(i = 1, \cdots, q)$ and $\theta_j = \theta_j(\theta_{q+1}', \cdots, \theta_l')$ $(j = q + 1, \cdots, l)$. Consequently,

$$\begin{aligned} p'(x_1, \cdots, x_n \mid \theta_1', \cdots, \theta_q', \cdots, \theta_l') \\ = p'(T_1, \cdots, T_m \mid \theta_1', \cdots, \theta_q')\cdot\phi'(x_1, \cdots, x_n ; \theta_{q+1}', \cdots, \theta_l') \end{aligned} \tag{4}$$

and the theorem is established.

NEW YORK, N. Y.

NOTE ON THE MOMENTS OF A BINOMIALLY DISTRIBUTED VARIATE

BY W. D. EVANS

J. A. Joseph, has given two interesting triangular arrangements of numbers, the second of which is reproduced herewith as Table 1.[1] The successive rows in this table are the coefficients in the expansion of x^n as a function of the factorials $x^{(i)}$, using the notation of the calculus of finite differences. For example,

$$x^4 = x^{(4)} + 6x^{(3)} + 7x^{(2)} + x,$$

where

$$x^{(i)} = x(x - 1)(x - 2) \cdots (x - i + 1).$$

Joseph points out that the coefficients may be used to generate the numbers of Laplace.

---

[1] J. A. Joseph, "On the Coefficients of the Expansion of $X^{(n)}$," *Annals of Math. Stat.*, Vol. X (1939), p. 293.

A general expression defining any of the coefficients in terms of its place of occurrence in Table 1 may be set up. If we denote by $F_c(r)$ the number in row r and column c of the table, we have

$$(1) \qquad F_c(r) = \sum_1^{r-c+1} k_1 \sum_1^{k_1} k_2 \sum_1^{k_2} k_3 \cdots \sum_1^{k_{c-2}} k_{c-1} \qquad (r \geq c).$$

This expression is of additional interest since the numbers defined by it are likewise the coefficients in the expression of the z-th moment about the origin of a binomially distributed variate in terms of the probability of the variate and the size of the sample in which it is contained. For example, it may be easily

TABLE 1

	1	2	3	4	5	$\cdots$	c
1	1						
2	1	1					
3	1	3	1				
4	1	6	7	1			
5	1	10	25	15	1		
:	:	:	:	:	:		
r	$F_1(r)$	$F_2(r)$	$F_3(r)$	$F_4(r)$	$F_5(r)$	$\cdots$	$F_c(r)$

verified that if α is such a variate, p its probability of occurrence, and n the size of the sample in which it is contained,

$$E(\alpha)^2 = n^{(2)}p^2 + np$$

$$E(\alpha)^3 = n^{(3)}p^3 + 3n^{(2)}p^2 + np$$

$$E(\alpha)^4 = n^{(4)}p^4 + 6n^{(3)}p^3 + 7n^{(2)}p^2 + np$$

and so on.

Ordinarily, computation of the higher moments of a binomially distributed variate is a tedious process of repeated differentiation. However, equation (1) immediately permits us to generalize the foregoing expressions to give the z-th moment of α as follows:

$$(2) \qquad E(\alpha)^z = \sum_{i=0}^{z-1} n^{(z-i)} p^{z-i} \sum_1^{z-i} k_1 \sum_1^{k_1} k_2 \cdots \sum_1^{k_{i-1}} k_i .$$

It will be noted that when $c - 1$ in equation (1) and i in equation (2) are equal to zero, the repeated summations vanish to be replaced by the value one.

By means of equation (2) much of the labor usually involved in expressing the z-th moment about the origin of a binomially distributed variate in terms of n and p may be avoided.

WASHINGTON, D. C.

REPORT OF THE ANNUAL MEETING OF THE INSTITUTE

The fifth annual meeting of the Institute of Mathematical Statistics was held in Philadelphia, Pennsylvania, on December 27 and 28, 1939, in conjunction with the meetings of the American Statistical Association, the Econometric Society, and the American Sociological Society. The program for the meeting was arranged by Professor C. C. Craig.

On Wednesday morning, December 27, the Institute held a session devoted to contributed papers on *Statistical Theory* and *Methodology*. Professor P. R. Rider, President of the Institute, presided. At that time the following papers were presented:

1. *On the unbiased character of certain likelihood-ratio tests when applied to normal systems.*
   Joseph F. Daly, The Catholic University of America.
2. *The product seminvariants of the mean and a central moment in samples.*
   C. C. Craig, University of Michigan.
3. *A method for minimizing the sum of absolute values of deviations.*
   Robert Singleton, Princeton Local Government Survey.
4. *On certain criteria for testing the homogeneity of k estimates of variance.*
   C. Eisenhart and Frieda S. Swed, University of Wisconsin.
5. *On a test whether two samples are from the same population.*
   A. Wald and J. Wolfowitz, Columbia University and Brooklyn, New York.
6. *The power functions of certain tests of significance in harmonic analysis and lag correlation.*
   William G. Madow, Washington, D. C.
7. *Some theoretical aspects of the use of transformations in the statistical analysis of replicated experiments.*
   W. G. Cochran, Iowa State College.
8. *The standard errors of geometric and harmonic types of index numbers.*
   Nilan Norris, Hunter College.
9. *A study of R. A. Fisher's z distribution and the related F distribution.*
   L. A. Aroian, Hunter College.
10. *A note on the analysis of variance with unequal class frequencies.*
    Abraham Wald, Columbia University.
11. *An approach to problems involving disproportionate frequencies.*
    Burton D. Seeley, U. S. Department of Labor.

Abstracts of these papers are given at the close of this report.

Immediately following the session just described, the Institute held its annual business meeting. At that time President Rider announced that the newly elected officers for the year 1940 are: President, S. S. Wilks, Princeton University; Vice-Presidents: C. C. Craig, University of Michigan, and A. T. Craig, University of Iowa; Secretary-Treasurer: P. R. Rider, Washington University.

At one o'clock on the same day, members of the Institute and their guests

attended the annual luncheon. At the luncheon, Professor B. H. Camp addressed the Institute on *Non-standard Deviations.*

On Wednesday afternoon, the Institute met jointly with the American Statistical Association for a program devoted to *Lag Effects in Statistics and Economics.* Professor J. D. Tamarkin presided and at this time the following papers were read:

1. *Lag effects in statistics and related problems.*
   A. J. Lotka, Metropolitan Life Insurance Company.
2. *Some methods in the analysis of lag effects.*
   H. T. Davis, Northwestern University.
3. *Lag effects in economics.*
   Charles F. Roos, Institute of Applied Econometrics, Inc.

A joint session with the Biometric Section of the American Statistical Association was held on Wednesday evening, Professor George W. Snedecor presiding. The papers presented at this session, which dealt with *Design and Analysis of Replicated Experiments,* were the following:

1. *Practical difficulties met in the use of experimental designs.*
   A. E. Brandt, Soil Conservation Service.
2. *Factorial design and covariance in the biological assay of vitamin D.*
   C. I. Bliss, Sandusky, Ohio.
3. *Combinatorial problems in the design of experiments.*
   Gertude M. Cox, Iowa State College.
4. *Experimental trials with balanced incomplete blocks.*
   W. J. Youden, Boyce Thompson Institute.

On Thursday afternoon the Institute held consecutively joint sessions with the American Sociological Society and the Econometric Society. At the first of these, Professor William F. Ogburn presided and the following program was presented:

1. *How the mathematician can help the sociologist.*
   Samuel A. Stouffer, University of Chicago.
2. *Some problems of combinations and permutations as they apply to a comprehensive classification of social groups.*
   George A. Lundberg, Bennington College.

Discussion: C. C. Craig, University of Michigan.
Philip M. Houser, U. S. Bureau of the Census.

At the second session the topic for discussion was *Recent Advances in Business Cycle Analysis* and these papers were given:

1. *Recursive methods in business cycle analysis.*
   Merrill M. Flood, Princeton Surveys.
2. *An appreciation of some recent mathematical business cycle theories.*
   Gerhard Tintner, Iowa State College.
3. *The statisticians' new clothiers.*
   Arne Fisher, Western Union Telegraph Company.

PAUL R. RIDER, *Secretary.*

ABSTRACTS OF PAPERS

(Presented on December 27, 1939, at the Philadelphia meeting of the Institute)

On the Unbiased Character of Certain Likelihood-Ratio Tests when Applied to Normal Systems. JOSEPH F. DALY, The Catholic University of America.

Consider a random sample of N observations on a set of variates $x^1, \cdots, x^q$, where $x^1, \cdots, x^k$ are assumed to be normally distributed about means which are linear functions $m^i = \Sigma b^i_\sigma x^\sigma$ of the fixed variates $x^{k+1}, \cdots, x^q$. One is sometimes required to decide whether the sample tends to contradict the further hypothesis, H_0, that the coefficients b^i_μ belonging to a certain subset of the fixed variates, say $x^{k+1}, \cdots, x^{k+h}$, have the specific values $b^i_{\mu 0}$. Such a situation occurs, for example, in the generalized analysis of variance. In this paper it is shown that the Neyman-Pearson method of the ratio of likelihoods yields a test of H_0 which is (at least locally) unbiased; in other words, this test is less likely to reject H_0 when the sample is in fact drawn from a normal population in which $b^i_\mu = b^i_{\mu 0}$ than when it is drawn from a normal population in which the b^i_μ are different from but sufficiently close to $b^i_{\mu 0}$. In the special cases $k = 1$ or $h = 1$ the proof goes through even without the restriction that the true b^i_μ be close to $b^i_{\mu 0}$, a result which is also implicit in the papers by P. C. Tang and P. L. Hsu (*Stat. Res. Mem.* Vol. 2).

Similarly with respect to the hypothesis H_I that the deviations $x^i - \Sigma b^i_\sigma x^\sigma$ fall into certain mutually independent sets the λ-test is at least locally unbiased; and it has the additional property that the expected value of any positive integral power of $\sqrt[N]{\lambda}$ is greater when H_I is true than when the sample is drawn from any other normal population.

The Product Seminvariants of the Mean and a Central Moment in Samples. C. C. CRAIG, The University of Michigan.

The method used by the author in calculating the product seminvariants of a pair of central moments in samples is not adapted without modification to the present problem. In the present paper the necessary modification is developed which gives a routine method for the calculation of these sampling distribution characteristics. The calculation is a little heavier than in the previous case but the results for the mean and the second, third, and fourth central moments are given up to the fourth order except in one case in which the weight is 13. It is planned to follow this with a further study of the distribution of Fisher's t in samples from a normal population.

A Method for Minimizing the Sum of Absolute Values of Deviations. ROBERT SINGLETON, Princeton Local Government Survey.

E. C. Rhodes (*Philosophical Magazine*, May 1930) presented a method for the estimation of parameters in a linear regression where it is desired to minimize the sum of absolute values of the deviations. In this paper the structure of the deviation surface is analyzed and a method of steepest descent is developed which for computational purposes is an improvement over Rhodes' method. The process is finite and leads to an exact solution. The method and the formulae used are such as to permit the successive additions of new observations or sets of observations to the original data, or the exclusion of an observation from the original set, and the determination of the parameters for the sets of data so derived, with little additional labor.

On Certain Criteria for Testing the Homogeneity of k Estimates of Variance. C. EISENHART AND FRIEDA S. SWED, University of Wisconsin.

Given k variance estimates $s_1^2, s_2^2, \cdots, s_k^2$ with $n_r s_r^2$, $(r = 1, 2, \cdots, k)$, independently distributed as $\chi^2\sigma_r^2$ for n_r degrees of freedom, tests of the hypothesis, H_0, that $\sigma_r^2 = \sigma^2$, $(r = 1, 2, \cdots, k)$, where σ^2 is unknown, have been based to date on one or the other of the quantities

$$Q_1 = \sum_{r=1}^{k} n_r(s_r^2 - s^2)^2/2s^4$$

$$Q_2 = w \log (ns^2/w) - \sum_{r=1}^{k} w_r \log \{n_r s_r^2/w_r\}$$

where the w_r are weights, $w = \sum_{r=1}^{k} w_r$, $n = \sum_{r=1}^{k} n_r$, and $ns^2 = \sum_{r=1}^{k} n_r s_r^2$. A. E. Brandt and W. L. Stevens have advocated the use of Q_1, referring an observed value of Q_1 to the χ^2 distribution for $k - 1$ degrees of freedom. J. Neyman, E. S. Pearson, B. L. Welch, and M. S. Bartlett have advocated tests based on Q_2, Bartlett definitely proposing the use of degrees of freedom as weights, i.e. $w_r = n_r$, and recent work of E. J. G. Pitman and others has shown that unless $w_r = n_r$ tests based on Q_2 are biased. (A statistical test of an hypothesis H is said to be unbiased when the probability of rejecting H by its use is a minimum when H is true; obviously a desirable property.) When $w_r = n_r$ Bartlett has suggested that the distribution of Q_2 can be satisfactorily approximated by referring $Q_2/\left\{1 + \frac{1}{3(k-1)} \cdot \left(\sum_{r=1}^{k} \frac{1}{n_r} - \frac{1}{n}\right)\right\}$ to the χ^2 distribution for $k - 1$ degrees of freedom. In this paper we discuss the adequacy of the χ^2 distribution to describe the distribution of Q_1 and of the adjusted Q_2 when the degrees of freedom, n_r, are small.

U. S. Nair and D. J. Bishop have given theoretical evidence which suggests that when $n_r \geq 2$, $(r = 1, 2, \cdots, k)$, Bartlett's adjusted Q_2 may be expected to conform to the χ^2 distribution reasonably well in the neighborhood of the 5% and 1% levels. Using 1000 samples of 4 for which $n_r s_r^2/(n_{r+1})$ has been tabulated by W. A. Shewhart in Table D, Appendix II of his "Economic Control of Quality of Manufactured Product," 200 values of Q_1 and Q_2 (with adjustment) were calculated and compared with the χ^2 distribution for $k - 1$ degrees of freedom. Two cases were studied: Case I, $k = 5$ and $n_1 = n_2 = \cdots = 3$; Case II, $k = 3$ and $n_1 = n_2 = 3$ while $n_3 = 9$. As measured by the Chi-Square Goodness of Fit Test, using 11 degrees of freedom, the fits were good in all four instances. In Case I, for Bartlett's adjusted Q_2 the test led to $.80 < P < .90$, and to $.70 < P < .80$ for the Brandt-Stevens Q_1; in Case II, the fits were poorer with $.50 < P < .70$ for Bartlett's criterion and $.10 < P < .20$ for the Brandt-Stevens. However, an examination of the *descending* cumulative distributions showed that in all instances these criteria exhibited a deficiency of large values of χ^2, with the deficiency, in general, more marked in the case of the Brandt-Stevens test. Consequently, when one uses significance levels for these criteria obtained by means of the χ^2 approximation advocated, one is in reality using a level of significance slightly less than that professed. The discrepancy is not great, however, and is on the safe side, i.e. one will reject H_0 falsely in the long run less often than one professes to be doing. Without doubt, however, one will also detect the falsehood of H_0 when $\sigma_r^2 \neq \sigma_t^2$, for at least one pair of values of r and t, $r \neq t$, less often in the long run by the use of these approximate significance levels than if the true levels were used, but we have no definite evidence at present on this point. A somewhat disquieting feature is that the agreement between the χ^2 values yielded by the two criteria becomes worse as one proceeds toward larger values of χ^2 in

terms of either quantity. Thus, of 8 samples which Q_2 would have rejected at the 5% level in Case I, only 4 of these would have been rejected by Q_1, and Q_2 would have passed 3 samples of the 7 rejected by Q_1. Thus it appears that, if one wishes to work with a given chance of rejecting H_0 falsely, one should choose one of these criteria and then stick to it in future applications. For large values of the n_r the two criteria tend to equivalence, so the choice between them is of interest mainly for small n_r, but cannot be made with full information until more is known about the bias, if any, of the Brandt-Stevens test, and the relative power of the two tests with regard to alternatives to H_0.

On a Test Whether Two Samples are from the Same Population. A. WALD AND J. WOLFOWITZ, Columbia University and Brooklyn, New York.

Let X and Y be two independent random variables about whose distributions nothing is known except that they are continuous. Let $x_1, x_2, \cdots, x_m$ be a set of m independent observations on X and let $y_1, y_2, \cdots, y_n$ be a set of n independent observations on Y. The null hypothesis to be tested is that the distributions of X and Y are identical.

Let the set of $m + n$ observations be arranged in order of magnitude, thus: $z_1, z_2, \cdots, z_{m+n}$. Replace z_i by v_i $(i = 1, 2, \cdots, m + n)$ where $v_i = 0$ if z_i is a member of the set of x's and $v_i = 1$, if z_i is a member of the set of y's. Since the null hypothesis states only that the distributions of X and Y are identical without specifying them in any other way, the distribution of the statistic U used for testing the null hypothesis must be independent of this common distribution of X and Y. It can easily be shown that the statistic U must be a function only of the sequence $v_1, v_2, \cdots, v_{m+n}$.

A subsequence $v_s, v_{s+1}, \cdots, v_{s+r}$ (where r may also be 0) is called a run if $v_s = v_{s+1} = \cdots = v_{s+r}$ and if $v_{s-1} \neq v_s$ when $s < 1$ and if $v_{s+r} \neq v_{s+r+1}$ when $s + r < m + n$. The statistic U defined as the number of runs in the sequence $v_1, v_2, \cdots, v_{m+n}$ seems a suitable statistic for testing the null hypothesis. A difference in the distribution functions of X and Y tends to decrease U. Hence the critical region is defined by the inequality $U < u_0$, where u_0 depends only on m, n, and the level of significance adopted. If $m \leq n$ and $P\{U = c\}$ is the probability that $U = c$, then:

$$P\{U = 2K\} = \frac{2({}^{m-1}C_{k-1} \cdot {}^{n-1}C_{k-1})}{{}^{m+n}C_m}, \qquad (K = 1, 2, \cdots, m),$$

$$P\{U = 2K - 1\} = \frac{({}^{m-1}C_{k-1}{}^{n-1}C_{k-2} + {}^{m-1}C_{k-2}{}^{n-1}C_{k-1})}{{}^{m+n}C_m}, \qquad (K = 2, 3, \cdots, m + 1).$$

The mean of U is:

$$\frac{2mn}{m + n} + 1.$$

The variance of U is:

$$\frac{2mn(2mn - m - n)}{(m + n)^2(m + n - 1)}.$$

If $\frac{m}{n} = \alpha$ (a positive constant) and $m \to \infty$, the distribution of U converges to the normal distribution.

The Distribution of Quadratic Forms In Non-Central Normal Random Variables. WILLIAM G. MADOW, Washington, D. C. (Presented to the Institute under a slightly different title)

Let the distribution of a sum of non-central squares of normally and independently distributed random variables which have the unit variances be called the χ'^2 distribution. It is proved that if a set of quadratic forms have a sum which is the sum of the squares of their variables, then a necessary and sufficient condition that the quadratic forms be independently distributed in χ'^2 distributions is that the rank of the sum of quadratic forms be equal to the sum of the ranks of the quadratic forms. Furthermore, the constants on which the χ'^2 distributions depend may be obtained by substituting the values about which the variables are taken for the variables themselves in the quadratic forms. Roughly speaking the theorem states that if a set of quadratic forms satisfy the conditions of the Fisher-Cochran theorem when the true means vanish, then the set of quadratic forms will be independently distributed in χ'^2 distributions when the true means do not vanish.

Some Theoretical Aspects of the Use of Transformations in the Statistical Analysis of Replicated Experiments. W. G. Cochran, Iowa State College.

The device of transforming the data to a different scale before performing an analysis of variance has recently been recommended by a number of writers for replicated experiments in which the original data show a markedly skew distribution. The use of transformations to obtain an approximate analysis has been supported mainly on the grounds that in the transformed scale the true experimental error variance is approximately the same on all plots. This paper considers the relation of the method of transformations to a more exact analysis. Discussion is confined to the $\sqrt{x}$ and $\sin^{-1}\sqrt{x}$ transformations, which appear to receive the most frequent use in practice.

To obtain an exact analysis, it is necessary to specify (i) how the expected value on any plot is obtained from unknown parameters representing the treatment and block (or row and column) effects (ii) how the observed values on the plots vary about the expected values. If the latter variation follows the Poisson law, (a case to which the square root transformation has been considered appropriate), the equations of estimation by maximum likelihood take the form

$$\sum_c \left(\frac{x - m}{m}\right)\frac{\partial m}{\partial c} = 0, \tag{1}$$

where x is the observed and m the expected value on any plot, c is a typical unknown parameter, and the summation extends over all plots whose expectations involve c. As the number of parameters is usually large (e.g. 16 in a 6 x 6 Latin square), these equations are laborious to solve; moreover, the question of obtaining small-sample tests of significance is difficult. It is shown that if a particular form can be assumed for the prediction formula in (i), namely that $\sqrt{m}$ is a linear function of the treatment and block (or row and column) constants, the equations of estimation may be reduced to the simpler form

$$\sum_c 4(r' - \sqrt{m}) = 0, \tag{2}$$

where $r' = \frac{1}{2}\left(\sqrt{m} + \frac{x}{\sqrt{m}}\right)$ is a function closely related to the square root of x. It follows that the statistical analysis in square roots, with some slight adjustments, coincides with the maximum likelihood solution, provided that the above form can be assumed for the prediction formula. The appropriateness of this form in practice is briefly considered and a "goodness of fit" test by χ^2 is developed. A numerical example is worked as an illustration and indicates that a good approximation is obtained by the transformation alone even with very small numbers per plot. The corresponding theory is also discussed for the inverse sine transformation, which applies where the original data are percentages or fractions whose experimental errors are derived from the binomial distribution.

In practice the type of analysis outlined above is unlikely to supplant the simple use of transformations, because it can seldom be assumed that the experimental variance is entirely of the Poisson or binomial type. The more exact analysis may, however, be useful (*i*) for cases in which the plot yields are very small integers or the ratios of very small integers (*ii*) in showing how to give proper weight to an occasional zero plot yield.

The Standard Errors of Geometric and Harmonic Types of Index Numbers. By Nilan Norris, Hunter College.

Various statisticians have made empirical studies of the sampling errors of certain types of index numbers used in the United States and England. None of these writers has taken advantage of the tools afforded by the modern theory of estimation, including fiducial inference, as a means of arriving at direct and general expressions for estimating the standard deviations of the sampling errors of geometric and harmonic types of index numbers.

A known expression for the first approximation to the variance of a function, as given by the relation between the variance of the function and the variance of the argument, is valid for that general class of distributions of which the variance and a higher moment are finite. With the aid of this relation, there appear simple and useful forms for estimating the standard errors of geometric and harmonic types of indexes. For sufficiently large samples, these forms are valid for all of the types of distributions of price relatives, production relatives, and similar observations ordinarily encountered, provided that there are satisfied the necessary conditions for drawing sound inferences on the basis of sampling without reference to the value of the variate.

Necessary conditions for using tests of significance soundly in connection with index number problems are those of realistic and intimate acquaintance with observations, and careful attention to certain broad theoretical considerations which determine whether or not the index is suited for the purpose for which it is used.

A Study of R. A. Fisher's z Distribution and the Related F Distribution. L. A. Aroian, Hunter College.

The following results for the z distribution and related F distribution are investigated:

(1) Geometric properties.

(2) Exact values of the seminvariants and moments of z. Exact values of the first four central moments of F.

(3) The approach to normality of both distributions as n_1 and n_2 become large in any manner whatever.

(4) The Pearson types of approximating curves, the logarithmic normal approximation, the Gram-Charlier approximation, and the uses of these in finding any level of significance of z and of F.

A Note on the Analysis of Variance with Unequal Class Frequencies. Abraham Wald, Columbia University.

Let us consider p groups of variates and denote by m_j ($j = 1, \cdots, p$) the number of elements in the j-th group. Let x_{ij} be the i-th element in the j-th group. We assume that x_{ij} is the sum of two variates ϵ_{ij} and η_j, i.e. $x_{ij} = \epsilon_{ij} + \eta_j$ where ϵ_{ij} ($i = 1, \cdots, m_j$; $j = 1, \cdots, p$) is normally distributed with mean μ and variance σ^2, and η_j ($j = 1, \cdots, p$) is normally distributed with mean μ' and variance σ'^2. All the variates ϵ_{ij} and η_j are supposed to be distributed independently. The intra-class correlation ρ is given by

$$\rho = \frac{\sigma'^2}{\sigma^2 + \sigma'^2}.$$

Confidence limits for ρ have been derived only in case of equal class frequencies, i.e. $m_1 = m_2 = \cdots = m_p$. We give here the confidence limits for ρ in case of unequal class frequencies. Since ρ is a monotonic function of $\frac{\sigma'^2}{\sigma^2}$, it is sufficient to derive confidence limits for $\frac{\sigma'^2}{\sigma^2}$. Denote $\frac{\sigma'^2}{\sigma^2}$ by λ^2 and the arithmetic mean of the j-th group by $\bar{x}_j$. Let

$$w_j = \frac{m_j}{1 + m_j\lambda^2},$$

and denote by F_1 and F_2 the lower and upper confidence limits respectively of F, where F has the analysis of variance distribution with $p - 1$ and $N - p = m_1 + \cdots + m_p - p$ degrees of freedom. Then the lower confidence limit λ_1^2 of λ^2 is given by the root of the equation in λ^2:

$$f(\lambda^2) = \frac{N - p}{p - 1} \cdot \frac{\sum_{j=1}^{p} \left\{ w_j \left(\bar{x}_j - \frac{\Sigma w_j \bar{x}_j}{\Sigma w_j} \right)^2 \right\}}{\Sigma\Sigma(x_{ij} - \bar{x}_j)^2} = F_2, \tag{1}$$

and the upper confidence limit λ_2^2 of λ^2 is given by the root of

$$f(\lambda^2) = F_1. \tag{2}$$

For calculating the roots of (1) and (2), we can make use of the fact that $f(\lambda^2)$ is monotonically decreasing with increasing λ^2.

An Approach to Problems Involving Disproportionate Frequencies. Burton D. Seeley, Washington, D. C.

Applied mechanics offers an analysis of variance solution to problems of multiple classification involving disproportionate sub-class numbers. The quality of orthogonality may be attained in such problems by measuring the variability between classes of any one classification after centering the others. This approach, which is not limited by the number of classes or the number of classifications, treats the problem involving equal sub-class numbers as a special phase of the general analysis of variance.

CONSTITUTION
OF THE
INSTITUTE OF MATHEMATICAL STATISTICS

ARTICLE I

Name and Purpose

1. This organization shall be known as the Institute of Mathematical Statistics.
2. Its object shall be to promote the interests of mathematical statistics.

ARTICLE II

Membership

1. The membership of the Institute shall consist of Members, Fellows, Honorary Members, and Sustaining Members.
2. Voting members of the Institute shall be (a) the Fellows, and (b) all others who have been members for twenty-three months prior to the date of voting.

ARTICLE III

Officers, Board of Directors, Committee on Membership, and Committee on Publications

1. The Officers of the Institute shall be a President, two Vice-Presidents, and a Secretary-Treasurer, elected for a term of one year by a majority ballot at the annual meeting of the Institute. Voting may be in person or by mail.

(a) Exception. The first group of Officers shall be elected by a majority vote of the individuals present at the organization meeting, and shall serve until December 31, 1936.

2. The Board of Directors of the Institute shall consist of the Officers and the previous President.
3. The Institute shall have a Committee on Membership composed of three Fellows. At their first meeting subsequent to the adoption of this Constitution, the Board of Directors shall elect three members as Fellows to serve as the Committee on Membership, one member of the Committee for a term of one year, another for a term of two years, and another for a term of three years. Thereafter the Board of Directors shall elect from among the Fellows one member annually at their first meeting after their election for a term of three years. The president shall designate one of the Vice-Presidents as Chairman of this Committee.
4. The Institute shall have a Committee on Publications composed of three Members or Fellows elected by the Board of Directors. The President shall designate a Vice-President as Ex Officio Chairman of this Committee.

ARTICLE IV

Meetings

1. A meeting for the presentation and discussion of papers, for the election of Officers, and for the transaction of other business of the Institute shall be held annually at such time as the Board of Directors may designate. Additional meetings may be called from

time to time by the Board of Directors and shall be called at any time by the President upon written request from ten Fellows. Notice of the time and place of meeting shall be given to the membership by the Secretary-Treasurer at least thirty days prior to the date set for the meeting. All meetings except executive sessions shall be open to the public. Only papers accepted by a Program Committee appointed by the President may be presented to the Institute.

2. The Board of Directors shall hold a meeting immediately after their election and again immediately before the expiration of their term. Other meetings of the Board may be held from time to time at the call of the President or any two members of the Board. Notice of each meeting of the Board, other than the two regular meetings, together with a statement of the business to be brought before the meeting, must be given to the members of the Board by the Secretary-Treasurer at least five days prior to the date set therefor. Should other business be passed upon, any member of the Board shall have the right to reopen the question at the next meeting.

3. The Committee on Membership shall hold a meeting immediately after the annual meeting of the Institute. Further meetings of the Committee may be held from time to time at the call of the Chairman or any member of the Committee provided notice of such call and the purpose of the meeting is given to the members of the Committee by the Secretary-Treasurer at least five days before the date set therefor. Should other business be passed upon, any member of the Committee shall have the right to reopen the question at the next meeting.

4. At a regularly convened meeting of the Board of Directors, three members shall constitute a quorum. At a regularly convened meeting of the Committee on Membership, two members shall constitute a quorum.

ARTICLE V

Publications

1. The *Annals of Mathematical Statistics* shall be the Official Journal for the Institute. Other publications may be originated by the Board of Directors as occasion arises.

ARTICLE VI

Expulsion or Suspension

1. Except for non-payment of dues, no one shall be expelled or suspended except by action of the Board of Directors with not more than one negative vote.

ARTICLE VII

Amendments

1. This constitution may be amended by an affirmative two-thirds vote at any regularly convened meeting of the Institute provided notice of such proposed amendment shall have been sent to each voting member by the Secretary-Treasurer at least thirty days before the date of the meeting at which the proposal is to be acted upon. Voting may be in person or by mail.

BY-LAWS

ARTICLE I

Duties of the Officers, Board of Directors, Committee on Membership, and Committee on Publications

1. The President, or in his absence, one of the Vice-Presidents, or in the absence of the President and both Vice-Presidents, a Fellow selected by vote of the Fellows present,

shall preside at the meetings of the Institute and of the Board of Directors. At meetings of the Institute, the presiding officer shall vote only in the case of a tie, but at meetings of the Board of Directors he may vote in all cases. At least three months before the date of the annual meeting, the President shall appoint a Nominating Committee of three members. It shall be the duty of the Nominating Committee to make nominations for Officers to be elected at the annual meeting and the Secretary-Treasurer shall notify all voting members at least thirty days before the annual meeting. Additional nominations may be submitted in writing, if signed by at least ten Fellows of the Institute, up to the time of the meeting.

2. The Secretary-Treasurer shall keep a full and accurate record of the proceedings at the meetings of the Institute and of the Board of Directors, send out calls for said meetings and, with the approval of the President and the Board, carry on the correspondence of the Institute. Subject to the direction of the Board, he shall have charge of the archives and other tangible and intangible property of the Institute. He shall send out calls for annual dues and acknowledge receipt of same; pay all bills approved by the President for expenditures authorized by the Board or the Institute; keep a detailed account of all receipts and expenditures, prepare a financial statement at the end of each year and present an abstract of the same at the annual meeting of the Institute after it has been audited by a Member or Fellow of the Institute appointed by the President as Auditor. The Auditor shall report to the President.

3. The Board of Directors shall have charge of the funds and of the affairs of the Institute, with the exception of those affairs specifically assigned to the President or to the Committee on Membership. The Board shall have authority to fill all vacancies ad interim, occurring among the Officers, Board of Directors, or in any of the Committees. The Board may appoint such other committees as may be required from time to time to carry on the affairs of the Institute.

4. The Committee on Membership shall prepare and make available through the Secretary-Treasurer an announcement indicating the qualifications requisite for the different grades of membership.

5. The Committee on Publications, under the general supervision of the Board of Directors, shall have charge of all matters connected with the publications of the Institute, and of all books, pamphlets, manuscripts and other literary or scientific material collected by the Institute. Once a year this Committee shall cause to be printed in the Official Journal the Constitution and By-Laws and a classified list of all the Members and Fellows of the Institute.

ARTICLE II

Dues

1. Members shall pay five dollars at the time of admission to membership and shall receive the full current volume of the Official Journal. Thereafter, Members shall pay five dollars annual dues. The annual dues of Fellows shall be five dollars. The annual dues of Sustaining Members shall be fifty dollars. Honorary Members shall be exempt from all dues.

2. Annual dues shall be payable on the first day of January of each year.

3. The annual dues of a Fellow or Member include a subscription to the Official Journal. The annual dues of a Sustaining Member include two subscriptions to the Official Journal.

4. It shall be the duty of the Secretary-Treasurer to notify by mail anyone whose dues

may be six months in arrears, and to accompany such notice by a copy of this Article. If such person fail to pay such dues within three months from the date of mailing such notice, the Secretary-Treasurer shall report the delinquent one to the Board of Directors, by whom the person's name may be stricken from the rolls and all privileges of membership withdrawn. Such person may, however, be re-instated by the Board of Directors upon payment of the arrears of dues.

ARTICLE III

Salaries

1. The Institute shall not pay a salary to any Officer, Director, or member of any committee.

ARTICLE IV

Amendments

1. These By-Laws may be amended in the same manner as the Constitution or by a majority vote at any regularly convened meeting of the Institute, if the proposed amendment has been previously approved by the Board of Directors.

DIRECTORY OF THE INSTITUTE OF MATHEMATICAL STATISTICS

(As of January 1, 1940)

Please notify the Secretary, P. R. Rider, Washington University, St. Louis, Mo., of any changes of addresses or errors in names, titles or addresses as printed.

Anderson, Mr. Paul H. 108 E. Chalmers St., Champaign, Ill.
Acerboni, Dr. Argentino V. Larroque 232 Banfield, Buenos Aires, Argentina.
Alter, Prof. Dinsmore. Director of Griffith Observatory, Los Angeles, Calif.
Arnold, Prof. Herbert E. Dept. of Math., Wesleyan Univ., Middletown, Conn.
Aroian, Prof. Leo. Dept. of Math., Hunter College, Lexington Ave. and E. 68th St., New York, N. Y.
Bacon, Prof. H. M. Box 1144, Stanford Univ., Calif.
Baker, Dr. G. A. Experiment Station, College of Agriculture, Univ. of Calif., Davis, Calif..
Barral-Souto, Dr. José. Cordoba 1459, Buenos Aires, Argentina.
Barrett, Mr. C. S. 3145 Maple Ave., Brookfield, Ill.
Beall, Mr. Geoffrey. 729 Queen St., Chatham, Ontario, Canada.
Been, Mr. Richard O. 32 Old Fisheries Bldg., U. S. Bureau of Agricultural Economics, Washington, D. C.
Bennett, Prof. A. A. Dept. of Math., Brown Univ., Providence, R. I.
Berkson, Dr. Joseph. Mayo Clinic, Rochester, Minn.
Bernstein, Prof. Felix. Dept. of Biometry, New York Univ., New York, N. Y..
Boley, Mr. Charles C. 305 Ceramics Bldg., Urbana, Ill.
Bridger, Mr. Clyde A. 1801 No. Eleventh St., Boise, Idaho.
Brooks, Mr. A. G. 5819 W. Erie St., Chicago, Ill.
Burgess, Dr. R. W. Western Electric Co., 195 Broadway, New York, N. Y.
Burr, Prof. Irving W. Antioch College, Yellow Springs, Ohio.
Bushey, Prof. J. Hobart. Dept. of Math., Hunter College, Park Ave. and 68th St., New York, N. Y.
Camp, Prof. B. H. Dept. of Math., Wesleyan Univ., Middletown, Conn..
Camp, Prof. C. C. Dept. of Math., Univ. of Nebraska, Lincoln, Neb.
Carlson, Mr. John L. 736 West St., Reno, Nev.
Carver, Prof. H. C. Dept. of Math., Univ. of Michigan, Ann Arbor, Mich..
Chapman, Mr. Roy A. Forest Service, U. S. Dept. of Agriculture, 1000 Masonic Temple, New Orleans, La.
Clark, Prof. Andrew G. 1012 Laporte Ave., Ft. Collins, Colo.
Cochran, Prof. W. G. Statistical Laboratory, Iowa State College, Ames, Iowa.
Coleman, Mr. E. P. Dept. of Math., Municipal Univ. of Omaha, Omaha, Neb.
Cox, Dr. Gerald J. Mellon Institute, Pittsburgh, Pa.
Craig, Prof. A. T. Dept. of Math., Univ. of Iowa, Iowa City, Iowa..
Craig, Prof. C. C. Dept. of Math., Univ. of Michigan, Ann Arbor, Mich..
Crathorne, Prof. A. R. Dept. of Math., Univ. of Illinois, Urbana, Ill..
Crowe, Prof. S. E. 137 University Drive, East Lansing, Mich.
Curtiss, Dr. J. H. Belle Ayre Apts., Ithaca, N. Y.
Dahmus, Mr. Maurice E. 203 So. Fourth, Champaign, Ill.

Daly, Dr. Joseph F. 719 Irving St., N.E., Washington, D. C.
Dantzig, Mr. George B. 2609 Fulton St., Berkeley, Calif.
DeLury, Dr. D. B. Dept. of Math., Univ. of Toronto, Toronto 5, Canada.
Deming, Dr. W. E. Bureau of the Census, U. S. Dept. of Commerce, Washington, D. C..
Dodd, Prof. E. L. Dept. of Math., Univ. of Texas, Austin, Texas..
Dodge, Mr. Harold F. Bell Telephone Lab., 463 West St., New York, N. Y.
Doob, Dr. J. L. Dept. of Math., Univ. of Illinois, Urbana, Ill..
Dressel, Dr. Paul L. 6126 S. Woodlawn Ave. Chicago, Ill.
Dunlap, Prof. Jack W. Catharine Strong Hall, Univ. of Rochester, Rochester, N. Y.
Dwyer, Prof. Paul S. 407 Camden Ct., Ann Arbor, Mich.
Edgett, Dr. G. L. Queen's Univ., Kingston, Ontario, Canada.
Eisenhart, Dr. Churchill. Dept. of Math., Univ. of Wisconsin, Madison, Wis.
Elston, Mr. James S. Travelers Insurance Co., Hartford, Conn.
Elting, Mr. John P. Director of Research, Kendall Mills, Paw Creek, N. C.
Elveback, Miss Mary L. Princeton Surveys, Princeton, N. J.
Ettinger, Mr. W. J. Edison General Electric Appliance Co., 5600 W. Taylor St., Chicago, Ill.
Evans, Prof. H. P. North Hall, Univ. of Wisconsin, Madison, Wis.
Feldman, Dr. H. M. Sodan High School, St. Louis, Mo.
Fertig, Dr. John W. Dept. of Biostatistics, School of Hygiene and Public Health, Johns Hopkins Univ., Baltimore, Md.
Fischer, Dr. C. H. Dept. of Math., Wayne Univ., Detroit, Mich.
Fisher, Prof. Irving. 460 Prospect St., New Haven, Conn..
Flood, Dr. M. M. Princeton Surveys, Princeton, N. J.
Foster, Mr. Ronald M. 122 E. Dudley Ave., Westfield, N. J.
Frankel, Mr. Lester R. 1445 Otis Place, N.W., Washington, D. C.
Freeman, Prof. H. A. Dept. of Economics, Massachusetts Inst. of Technology, Cambridge, Mass.
Fry, Dr. T. C. Bell Telephone Lab., 463 West St., New York, N. Y..
Gavett, Prof. G. I. Dept. of Math., Univ. of Washington, Seattle, Wash.
Geiringer, Dr. Hilda. Bryn Mawr College, Bryn Mawr, Pa.
Gibson, Mr. Robert W. 503 W. California, Urbana, Ill.
Gill, Mr. John P. Dept. of Math., Univ. of Alabama, University, Ala.
Girshick, Mr. M. A. Bureau of Home Economics, Dept. of Agriculture, Washington, D. C.
Glover, Prof. J. W. 620 Oxford Rd., Ann Arbor, Mich..
Greenwood, Prof. J. A. College Station, Box 245, Durham, N. C.
Greville, Dr. Thos. N. E. Dept. of Math., Univ. of Michigan, Ann Arbor, Mich.
Grove, Prof. C. C. 143 Milburn Ave., Baldwin, L. I., N. Y.
Hart, Prof. W. L. Dept. of Math., Univ. of Minnesota, Minneapolis, Minn.
Hebley, Mr. Henry F. Product Control Manager, Pittsburgh Coal Co., P. O. Box 145, Pittsburgh, Pa.
Heidingsfield, Mr. Myron. Hotel Marcy, 720 West End Ave., New York, N. Y.
Henderson, Mr. Robert. Crown Point, Essex Co., N. Y..
Hendricks, Mr. Walter A. Division of Agricultural Statistics, Agricultural Marketing Service, U. S. Dept. of Agriculture, Washington, D. C.
Henry, Mr. M. H. Route No. 4, Box 394 M, Lansing, Mich.
Hoel, Dr. Paul G. Dept. of Math., Univ. of California at Los Angeles, Los Angeles, Calif.
Horst, Dr. Paul. Supervisor of Selection Research, Procter and Gamble, Cincinnati, Ohio.
Hotelling, Prof. Harold. Fayerweather Hall, Columbia Univ., New York, N. Y..
Huntington, Prof. E. V. 48 Highland St., Cambridge, Mass..
Hurwitz, Mr. William. 119 Concord Ave., Washington, D. C.
Ingraham, Prof. M. H. North Hall, Univ. of Wisconsin, Madison, Wis..
Jackson, Prof. Dunham. 119 Folwell Hall, Univ. of Minnesota, Minneapolis, Minn..

Jacobs, Mr. Walter. 5119 Third St., N.W., Washington, D. C.
Jankt, Dr. Jarslav. Praha XII, Krkonosska 3, Czechoslovakia.
Juran, Mr. J. M. Western Electric Co., 195 Broadway, New York, N. Y.
Keeping, Prof. E. S. Univ. of Alberta, Edmonton, Alberta, Canada.
Kelley, Prof. Truman L. Lawrence Hall, Cambridge, Mass..
Kenney, Mr. John F. Dept. of Math., Northwestern Univ., Evanston, Ill.
Kiernan, Prof. Charles J. 14 Fairbanks St., Hillside, N. J.
Knowler, Dr. Lloyd A. Dept. of Math., Univ. of Iowa, Iowa City, Iowa.
Knudsen, Miss Lila F. Junior Mathematician, Food Division, U. S. Dept. of Agriculture, Washington, D. C.
Kohl, Miss Alma. 2309 Russell St., Berkeley, Calif.
Kossack, Mr. Carl F. Dept. of Math., Univ. of Oregon, Eugene, Oregon.
Kullback, Dr. Solomon. Office of Chief Signal Officer, U. S. Navy Department, Washington, D. C..
Kurtz, Dr. Albert K. 173 Cornwall St., Hartford, Conn.
Laderman, Mr. Jack. 3224 Bronx Blvd., Bronx, N. Y.
Larsen, Prof. Harold D. Dept. of Math., Univ. of New Mexico, Albuquerque, N. M.
Leavens, Mr. Dickson H. Cowles Commission, Univ. of Chicago, Chicago, Ill.
Lemme, Prof. Maurice M. Dept. of Math., Univ. of Toledo, Toledo, Ohio.
Lindsey, Mr. Fred D. 2051 No. Brandywine St., Arlington, Va.
Livers, Mr. Joe J. Dept. of Math., Montana State College, Bozeman, Mont.
Lorge, Dr. Irving. 525 W. 120 St., New York, N. Y.
Lotka, Dr. A. J. One Madison Ave., New York, N. Y..
Lundberg, Prof. George A. Dept. of Sociology, Bennington College, Bennington, Vt.
Madow, Mr. William G. Hotel Lafayette, Washington, D. C..
Malzberg, Dr. Benjamin. New York State Dept. of Mental Hygiene, Albany, N. Y.
Mansfield, Prof. Ralph. Dept. of Math., Chicago Teachers College, 6800 Stewart Ave.. Chicago, Ill.
Mauchly, Prof. J. W. Dept. of Physics, Ursinus College, Collegeville, Pa.
Mayer, Mr. George F. T. 1519 Seventh St., S.E., Minneapolis, Minn.
Mises, Prof. R. von. Harvard Univ., Cambridge, Mass..
Mode, Prof. E. B. 688 Boylston St., Boston, Mass.
Molina, Mr. E. C. 463 West St., New York, N. Y..
Mudgett, Prof. Bruce D. Dept. of Economics, Univ. of Minnesota, Minneapolis, Minn.
MacLean, Mr. M. C. Chief of Census Analysis, Dominion Bureau of Statistics, Ottawa, Ontario, Canada.
McEwen, Prof. G. F. Scripps Institution, La Jolla, Calif.
Macphail, Dr. M. S. Dept. of Math., Acadia Univ., Wolfville, Nova Scotia, Canada.
McCarthy, Dr. Michael C. Dept. of Math., University College, Cork, Ireland.
Nesbitt, Dr. C. J. Dept. of Math., Univ. of Michigan, Ann Arbor, Mich.
Neyman, Prof. J. Dept. of Math., Univ. of California, Berkeley, Calif..
Norris, Dr. Nilan. Dept. of Economics, Hunter College, New York, N. Y.
Norton, Mr. K. A. 2123 Tunlaw Rd., N.W., Washington, D. C.
Olds, Prof. E. G. 953 La Clair Ave., Regent Square, Pittsburgh, Pa..
Ollivier, Dr. Arthur. Box 405, State College, Miss.
Olmstead, Dr. Paul S. 463 West St., New York, N. Y.
O'Toole, Dr. A. L. 6363 Sheridan Rd., Chicago, Ill..
Ozanne, Mr. Paul. Dept. of Agronomy, College of Agriculture, Univ. of Wisconsin, Madison, Wis.
Ore, Prof. Oystein. Dept. of Math., Yale Univ., New Haven, Conn.
Palley, Mr. A. 286 Commonwealth, Buffalo, N. Y.
Parente, Mr. A. R. 126 Church St., Hamden, Conn.
Perryman, Mr. James H. 12 Richmond Ave., Ilfracombe, Devon, England.

Petrie, Prof. George W. South Dakota School of Mines, Rapid City, S. Dak.
Pierce, Prof. J. A. Dept. of Math., Atlanta Univ., Atlanta, Ga.
Pixley, Prof. H. H. Dept. of Math., Wayne Univ., Detroit, Mich.
Pollard, Prof. H. S. Dept. of Math., Miami Univ., Oxford, Ohio.
Reed, Prof. L. J. 615 No. Wolfe St., Baltimore, Md.
Rider, Prof. P. R. Dept. of Math., Washington Univ., St. Louis, Mo..
Richardson, Prof. C. H. Bucknell Univ., Lewisburg, Pa.
Regan, Prof. Francis. Dept. of Math., St. Louis Univ., Grand and Pine Blvd., St. Louis, Mo.
Rietz, Prof. H. L. Dept. of Math., Univ. of Iowa, Iowa City, Iowa..
Romig, Mr. H. G. Bell Telephone Lab., 463 West St., New York, N. Y.
Roos, Dr. Charles F. Inst. of Applied Econometrics, 420 Lexington Ave., New York, N. Y..
Rulon, Prof. P. J. 13 Kirkland St., Cambridge, Mass.
Scarborough, Prof. J. B. P. O. Box 86, Annapolis, Md.
Schwartz, Mr. Herman M. College of Arts and Sciences, Duquesne Univ., Pittsburgh, Pa.
Shaw, Mr. Lawrence W. Fayerweather Hall, Columbia Univ , New York, N. Y.
Shewhart, Dr. W. A. 158 Lake Drive, Mountain Lakes, N. J..
Shohat, Prof. J. Dept. of Math., Univ. of Pennsylvania, Philadelphia, Pa..
Simon, Capt. Leslie E. Aberdeen Proving Ground, Aberdeen, Md.
Singleton, Mr. Robert R. 20 Nassau St., Princeton, N. J.
Spiegelman, Mr. Mortimer. 325 W. 86th St., New York, N. Y.
Stephan, Mr. F. F. 1626 K St., N. W., Washington, D. C.
Stouffer, Prof. S. A. Dept. of Sociology, Univ. of Chicago, Chicago, Ill.
Swanson, Mr. A. G. General Motors Inst., Flint, Mich.
Tai, Mr. Shih-Kuang. Tsing Hua Univ., Kun Ming, Yun Nan, China.
Thompson, Dr. William R. 27 Oakwood St., Albany, N. Y.
Thurstone, Prof. L. L. Dept. of Psychology, Univ. of Chicago, Chicago, Ill.
Toops, Prof. Herbert A. 458 W. 8th Ave., Columbus, Ohio.
Torrey, Miss Mary N. 463 West St., New York, N. Y.
Treloar, Prof. Alan E. Dept. of Biometry, Univ. of Minnesota, Minneapolis, Minn.
Trimble, Mrs. Anne. Apt. 612, 6104 Woodlawn Ave., Chicago, Ill.
Trubridge, Dr. G. F. P. 98 Kingsway, Petts Wood, Kent, England.
Vatnsdal, Mr. J. R. 304 Columbia St., Pullman, Wash.
Vickery, Dr. C. W. 304 E. 5th St., Austin, Texas.
Waite, Dr. Warren C. Division of Agricultural Economics, Univ. Farm, St. Paul, Minn.
Walker, Prof. Helen M. Teachers College, Columbia Univ., New York, N. Y.
Wald, Dr. A. Fayerweather Hall, Columbia Univ., New York, N. Y..
Wareham, Mr. Ralph E. 118 Elmer Ave., Schenectady, N. Y.
Wei, Dzung-shu. International House, 19 Evans St., Iowa City, Iowa.
Weida, Prof. F. M. Dept. of Math., George Washington Univ., Washington, D. C..
Welker, Mr. E. L. 160 Mathematics Bldg., Univ. of Illinois, Urbana, Ill.
Wescott, Mr. Mason E. 1936 Greenwood Ave., Wilmette, Ill.
White, Prof. A. E. Dept. of Math., Kansas State College, Manhattan, Kan.
Wilder, Dr. Marian. 115 Bedford S. E., Minneapolis, Minn.
Wilks, Prof. S. S. Fine Hall, Princeton Univ., Princeton, N. J..
Wilson, Dr. Elizabeth. One Waterman St., Cambridge, Mass.
Wilson, Mr. W. P. P. O. Box 1939, University, Ala.
Wolfenden, Mr. Hugh H. 182 Rosedale Heights Drive, Toronto 5, Canada.
Wright, Prof. Sewall. Dept. of Zoology, Univ. of Chicago, Chicago, Ill..
Wyckoff, Mr. J. F. Dept. of Math., Trinity College, Hartford, Conn.
Zoch, Mr. Richmond T. Cosmos Club, 1520 H. St., Washington, D. C.
Zubin, Dr. Joseph. 722 W. 168th St., New York, N. Y.